Reconstruction of the Last Eukaryotic Common Ancestor by cladistic and phylogenetic approach

Skejo, Josip

University of Zagreb
Faculty of Science

Josip Skejo

RECONSTRUCTION OF THE LAST EUKARYOTIC COMMON ANCESTOR BY CLADISTIC AND PHYLOGENETIC APPROACH

DOCTORAL THESIS

Zagreb, 2022.

University of Zagreb
Faculty of Science

Josip Skejo

RECONSTRUCTION OF THE LAST EUKARYOTIC COMMON ANCESTOR BY CLADISTIC AND PHYLOGENETIC APPROACH

DOCTORAL THESIS

Supervisors:
Prof William F. Martin
Assoc Prof Damjan Franjević

Zagreb, 2022.

Sveučilište u Zagrebu
Prirodoslovno – matematički fakultet

Josip Skejo

REKONSTRUKCIJA ZAJEDNIČKOGA PRETKA EUKARIOTA KLADISTIČKIM I FILOGENETIČKIM PRISTUPOM

DOKTORSKI RAD

Mentori:
Prof. dr. sc. William F. Martin
Izv. prof. dr. sc. Damjan Franjević

Zagreb, 2022.

The thesis was completed at the Institute of Molecular Evolution (Institut für Molekulare Evolution) of the Heinrich-Heine-Universität Düsseldorf, in Düsseldorf, Germany, under supervision of Professor Dr William F Martin and at the Evolution Lab of the Division of Zoology, Department of Biology, Faculty of Science, University of Zagreb in Croatia, under supervision of Associate Professor Dr Damjan Franjević. The research covered in this dissertation was funded by European Research Council (Grant No. 666053), the Volkswagen Foundation (Grant No. 93 046), and the Moore Simons Initiative on the Origin of the Eukaryotic Cell (Grant No. 9743).

The thesis is submitted to the University of Zagreb for the fulfilment of the doctoral study in biology requirements for the degree of Doctor of Science.

About the supervisors

Associate Professor **Dr Damjan Franjević**, born in Zagreb, Croatia, is an eminent and popular Croatian evolutionary biologist who works at the Division of Zoology of the Department of Biology, Faculty of Science, University of Zagreb, where he leads Evolution Lab and gives lectures in Evolution and Molecular Evolution.

Diploma, University of Zagreb [1997],
Master of Science in Molecular Biology, University of Zagreb [2002],
Doctor of Science in Biology, University of Zagreb [2006],
Supervised more than 70 bachelor and master theses [2006),
Published more than 40 scientific papers [1998),
Assistant professor at University of Zagreb [2016),
The Croatian National Science Award [2016],
Papers cited 715 times, h-index 13 [Google Scholar, 22.02.2022],
Husband and a father of three.

Professor **Dr William F Martin**, born in Bethesda, Maryland, USA, is a world-renowned evolutionary biologist, microbiologist, and botanist, with interest in biochemistry and physiology. Currently, he is the head of the Institute for Molecular Evolution at the Heinrich Heine Universität, Düsseldorf, Germany, where he also gives lectures in Genome and Proteome Analysis, Programming for Biologists, and Evolution and Biochemistry of Organelles.

Diploma, Technische Universität Hannover [1985],
PhD, Max-Planck-Institut für Züchtungsforschung, Cologne [1988],
Published more than 300 scientific papers [1986),
Heinz Maier-Leibnitz-Preis of the Deutsche Forschungsgemeinschaft [1990],
Habilitation, Technische Universität Braunschweig [1992],
Supervised hundreds of bachelors, master, and doctoral students [1992),
Full time professor at University of Düsseldorf [1994),
Technology Transfer Prize, Industrie und Handelskammer Braunschweig [1997],
Miescher-Ishida Prize of the International Society of Endocytobiology [1998],
Spiridion Brusina Medal of the Croatian Society of Natural Sciences [2017],
Preis der Klüh Stiftung [2018],
Papers cited 40922 times, h-index 97 [Google Scholar, 22.02.2022],
Editor of many scientific journals,
Husband and a father of four.

Acknowledgements

Many people are to be thanked why I am sitting here and writing this thesis, but I will try to keep it short. Thank you to post-docs who supervised my work and did numerous analyses with me, Sriram G Garg, Fernando DK Tria; thank you to Joana Xavier for introducing me to the promiscuous world of prokaryotes; thank you to Sven B Gould for sparking my interest in getting knowledge of the photosynthetic world; thank you to my students Karmela Adžić, Nico Bremer, Maks Deranja, Niko Kasalo, Ana Kulessa, Marko Pavlović, Jan Regul and Ivan Šapina for keeping me active and enthusiastic; thank you to Michael Hendriksen for all the help and discussion on mathematics; thank you to all the other co-workers, friends and co-authors whom I have published or discussed with: Neil Blackstone, Julia Brueckner, Niel Kapust, Michael Knopp, Falk Nagies, Jessica L Wimmer, Verena Zimorski. Finally, many thanks to my supervisors Damjan Franjević, and William F Martin. I am aware how difficult I am, so I can only thank you for what you have endured those five years with me. Fala! Thank you to all of you, once again, for your data, time and energy!

I dedicate the thesis to my teachers and my students.

BASIC DOCUMENTATION CARD

University of Zagreb
Faculty of Science
Department of Biology

Doctoral thesis

Reconstruction of the Last Eukaryotic Common Ancestor by cladistic and phylogenetic approach

JOSIP SKEJO

Organisation and diversity of the eukaryotic cells are well documented phenomena today, so it is already known scientific theory that LECA (the Last Eukaryotic Common Ancestor) harboured mitochondrion, nucleus, endoplasmic reticulum made of bacterial lipids, sex, meiosis, and eukaryotic life cycle. None of the aforementioned traits has been found in prokaryotes. The origin of those eukaryotes traits has, on the other side, been an unresolved issue for many years. It is questionable whether the mitochondrion (bacterial endosymbiont) entered the archaeal host cell prior to the formation of the nucleus or later; it has never been systematically studied if LECA had a single or many nuclei; it has not been quantified how many gene duplications there were in LECA, and what is their origin. Furthermore, cladistic terminology is often misapplied in molecular evolution and studies are rarely based on taxon-rich sampling, meaning that eukaryotic diversity is not well covered. This dissertation, hence, aims to discuss cladistics and its interpretation in theory, with emphasis on the definition of eukaryotes; and practically, by analyses, to shed some light on the genome organization, morphology, and physiology of LECA. Eukaryotes/Eukarya are defined as a monophyletic, holophyletic group with polyphyletic, reticulated origin. Because of the reticulated origin of eukaryotes and because mitochondria and the nucleus are not regarded as prokaryotes anymore, both Bacteria and Archaea are paraphyletic, i.e., monophyletic groups. Many duplications were present in LECA and the Bacteria-derived ones were found to be prevalent what suggests that mitochondria-early hypothesis might be correct. Except for the genes that originated from plastid acquisition (Cyanobacteria), no specific genes were found within eukaryotic supergroups, suggesting that differential loss and genome duplications are the major forces of the eukaryotic evolution. Ubiquity of the multinucleate state across the eukaryotic domain is presented. Traits annotated on the eukaryotic tree were multinucleate state presence and absence, open vs. closed nuclear division, as well as 'control traits' for which it is known to be ancestral to LECA (presence of sex, mitochondria) or for which it is known not to be ancestral to LECA (plastid, polyploidy). Ancestral state reconstruction did not reject the hypothesis that LECA was multinucleated, similar to modern aseptate fungi or myxomycetes, and exhibited closed nuclear division. It is confirmed (i.e., not rejected) that LECA was sexual, had mitochondria, did not have plastid and was not polyploid. The results of ancestral state investigations presented in this work indicate that, contrary to popular beliefs, LECA was likely not a uninucleate cell, from which it follows that uninuclear eukaryotes possibly represent highly specialized forms, of which some, such as Excavata might even have originated long time ago from LECA's gametes.

(146 pages, 19 figures, 278 references, three papers, original in English)

The thesis has been deposited at the Central Biological Library (Department of Biology, Faculty of Science, University of Zagreb, Marulićev trg 20/II, 10000 Zagreb, Croatia).

Keywords: evolution, eukaryogenesis, LECA, gametes, duplications, syncytium, coenocyte

Supervisors: Associate Professor Damjan Franjević, *University of Zagreb, Faculty of Science, Croatia*
Prof William F Martin, *Heinrich-Heine-Universität Düsseldorf, Germany*

Reviewers: Professor Jasna Hrenović, *University of Zagreb, Faculty of Science (Croatia)*
Professor Marija Heffer, *University of Osijek, Faculty of Medicine (Croatia)*
Professor Ivan Mijaković, *Technical University of Denmark (Denmark)*, *Chalmers University of Technology (Sweden)*
Professor Kristian Vlahoviček, *University of Zagreb, Faculty of Science (Croatia)*

Thesis accepted

TEMELJNA DOKUMENTACIJSKA KARTICA

Sveučilište u Zagrebu
Prirodoslovno – matematički fakultet
Biološki odsjek

Doktorska disertacija

Rekonstrukcija zajedničkoga pretka eukariota kladističkim i filogenetičkim pristupom

JOSIP SKEJO

Raznolikost eukariota i organizacija njihovih stanica danas su već dobro poznati pa se može smatrati znanstvenom teorijom da je LECA (engl. *Last Eukaryotic Common Ancestor*, Posljednji zajednički predak eukariota) imao je mitohondrij, jezgru, unutarstaničnu mrežu membrana građenu od bakterijskih lipida, spol, mejozu i eukariotski životni ciklus. U prokariota, usporedbe radi, ne postoji ni jedno od ovih obilježja. Usprkos dobrom poznavanju njihove građe, postanak eukariota i dalje je jedan od neriješenih problema moderne znanosti. Ne zna se je li mitohondrij (bakterijski endosimbiont) ušao u arheju prije ili nakon formacije jezgre; nikad nije utvrđeno je li LECA imao samo jednu ili je imao više jezgara; do sada nije određeno koliko je LECA imao dupliciranih gena i koje im je porijeklo. Uza sve to, u molekularnoj evoluciji često se pogrešno koristi terminologija sistematike, a radovi najčešće ne uključuju veliku raznolikost, tj. nije pokriveno mnogo svojti. Ciljevi ove disertacije jesu teoretski raspraviti terminologiju sistematike i njezinu interpretaciju, s naglaskom na definiciju eukariota; kao i istražiti organizaciju genoma, morfologiju i fiziologiju LECA-e praktičnim analizama. Eukarioti (Eukaryota/Eukarya) su definirani kao monofiletska; holofiletska skupina polifiletskog postanka. Zbog hibridnog postanka eukariota i budući da se mitohondrij i jezgra više ne smatraju prokariotima, i bakterije (Bacteria) i arheje (Archaea) su parafiletske tj. monofiletske skupine. Otkriveno je da je LECA imala mnogo duplikacija i da je najviše onih koje potječu od bakterija što ide u prilog hipotezi da je mitohondrij bio prisutan tijekom same eukariogeneze. Osim plastidnih gena koji su porijeklom iz cijanobakterija (*Cyanobacteria*), nisu pronađeni jedinstveni geni niti u jednoj supergrupi što znači da su veliki gubitci dijelova genoma i duplikacije gena glavne sile evolucije eukariotskog genoma. Višejezgrene stanice su učestale među eukariotima i u disertaciji je prikazana njihova raznolikost. Na stablu eukariota označene su prisutnost i odsutnost višejezgrenih stanica te otvorena i zatvorena dioba jezgre, kao i „kontrolna svojstva“, tj. ona za koja se zna da li su predačka LECA-i (spol, mitohondrij) i ona za koja se zna da nisu (plastid, poliploidija). Rekonstrukcija predačkih svojstava nije odbacila pretpostavke da je LECA bila višejezgrena, vjerojatno slična današnjim aseptičnim gljivama i sluznjačama (*Myxomycetes*), kao ni da je imala zatvorenu diobu jezgri. Potvrđeno je također, tj. nije odbačeno, da je LECA bila spolni organizam, da je imala mitohondrije, da nije imala plastide i da nije bila poliploidna. Suprotno uvriježenom mišljenju, LECA najvjerojatnije nije bila jednojezgreni organizam i prema tome jednojezgreni jednostanični eukarioti vjerojatno predstavljaju visokospecijalizirane organizme, od kojih neki poput supergrupe Excavata—smisleno nagađanje—možda čak potječu od LECA-inih gameta.

(146 stranica, 19 slika, 278 referenci, tri rada, original na engleskom)

Disertacija je pohranjena u Središnjoj biološkoj knjižnici Biološkog odsjeka Prirodoslovno – matematičkog fakulteta sveučilišta u Zagrebu (Marulićev trg 20/II, 10000 Zagreb, Hrvatska).

Ključne riječi: evolucija, eukariogeneza, LECA, gamete, duplikacije, sincicij, cenocit

Mentori: izv. prof. dr. sc. Damjan Franjević, *Sveučilište u Zagrebu, Prirodoslovno-matematički fakultet, Hrvatska*
prof. dr. sc. William F. Martin, *Heinrich-Heine-Universität Düsseldorf, Njemačka*

Ocjenjivači: prof. dr. sc. Jasna Hrenović, *Sveučilište u Zagrebu, Prirodoslovno-matematički fakultet (Hrvatska)*
prof. dr. sc. Marija Heffer, *Sveučilite u Osijeku, Medicinski fakultet, (Hrvatska)*
prof. dr. sc. Ivan Mijaković, *Technical University of Denmark (Danska)*, *Chalmers University of Technology (Švedska)*
prof. dr. sc. Kristian Vlahoviček, *Sveučilište u Zagrebu, Prirodoslovno-matematički fakultet (Hrvatska)*

Rad prihvaćen

Table of contents

Table of figures

INTRODUCTION

„Nothing makes sense, except in the light of evolution. “

T. Dobzhansky (1973)

Eukaryogenesis represents the most peculiar **symbiogenesis** in the evolutionary history by the means of which the first **eukaryote** originated (Mereschkowsky [Мережковский] 1910; Aanen & Eggleton 2017; Kowallik & Martin 2021). Parts of eukaryogenesis are, today, well understood, but there is much more information that is yet to be discovered. From the investigation of the structure of the members of the extant eukaryotic lineages, much information was deduced in a recent few decades (Thiergart et al. 2012; Ku et al. 2015a; Roger et al. 2017; Nagies et al. 2020) about the Last Eukaryotic Common Ancestor, the LECA (Martin & Müller 1998, Martin et al. 2015; O'Malley et al. 2019). For example, it is now undebated that it had **mitochondria** that originated from proteobacteria (Fan et al. 2020), that it had **endoplasmic reticulum** and **nucleus** that likely originated from bacterial vesicles (Gould et al. 2016; Gill et al. 2019.), as well as that its **genome** represents a **chimera** (Ku et al. 2015b, Brückner & Martin 2020, Knopp et al. 2021) between an archaeon and a bacterium (Imachi et al. 2020). The origin of eukaryotes occurred about two billion years ago (Radzvilavicius & Blackstone 2015, Porter 2020, Roger et al. 2021), and its result, the eukaryotes, have populated all habitats and ecological niches ever since (Mora et al. 2011, Adl et al. 2007, 2012). The same amount of time had passed from the origin of life to eukaryotes, as had passed between the eukaryotes and today (Booth & Doolittle 2015, Porter 2020), and because of that, it is important for one to understand the pre-eukaryotic world, i.e., the world of prokaryotes, in order to understand which events led to the rise of eukaryotes. The world we observe today is dominated by eukaryotic forms, the most famous being plants, animals, and fungi (Adl et al. 2012, Cavalier-Smith 1981, 1998, Nielsen 2012), so it is not easy to imagine there was a period on Earth when everything looked completely different (Martin & Russel 2003, Martin et al. 2008, Ferus et al. 2017, Weiss et al. 2018).

This thesis is a small coloured stone piece aiming to contribute to the whole mosaic of eukaryotic-origin research. In order to address questions about early eukaryotes and eukaryogenesis in their Earth history context and their scientific historical context, this introduction provides a brief insight into the origin of life, the diversity of prokaryotes, and various hypotheses aiming to explain eukaryogenesis. Finally, a brief overview of all the major eukaryotic lineages is presented (Adl et al. 2019), including schematic explanations of the primary and secondary plastids' origins (Gould et al. 2008). It is important to understand the modern eukaryotic diversity to its full extent in order to answer questions about the origin of this very diversity, and, of course, in order to become familiar with questions that are not currently answerable.

ORIGIN OF LIFE

Life began some 4-3.5 billion years ago in the deep ocean (Arndt & Nisbet 2012), and most probably at the **hydrothermal vents** (Baross & Hoffmann 1985; Russelll & Gall 1997; Martin et al. 2008). The Last Universal Common Ancestor (**LUCA**) marks the moment in which a part of Earth's **geochemistry** gave rise to the Life's **biochemistry** (Martin et al. 2014, Preiner et al. 2018). Traces of this ancient geochemistry are still found in biochemistry, and example are the iron–sulphur (FeS) clusters (Weiss et al. 2016). For a long time, origin of the life on Earth was explained via ***prebiotic organic soup*** (Haldane 1929, Опарин [Oparin] 1941), meaning that the early ocean was so full of organic compounds that redox reactions leading to the origin of life started spontaneously (e.g., Vincent et al. 2021). In such a historical concept, one could suppose that, for a moment, the whole **ocean** was LUCA, because the prebiotic reactions had been scattered everywhere (Martin et al. 2008). When the hydrothermal vents were discovered in a deep ocean in the 20th and the beginning of the 21th century (Corliss et al. 1979, Kelley et al. 2001), there was such a huge brand-new spectrum of the extant geochemical reactions discovered (Albarède 2009; Chester 2009), that it was the only logical approach for the whole field of the origin-of-life research to be turned upside down (Martin et al. 2008). With better understanding of the modern geochemical processes, it was much easier to find connections between abiotic and biotic chemistry (Martin & Russell 2003, Hazen & Sverjensky 2010).

Modern concepts of the LUCA are very different from the organic soup concept (Martin & Russell 2003). Where sea currents (rich in iron and carbon dioxide) and hydrothermal currents (rich in hydrogen) penetrate into pyrite rocks is the place where early **redox** started and where the transition between the geochemistry and the biochemistry occurred (Figure 1) (Koonin & Martin 2005; Martin et al. 2008): from simple redox reactions *via* the DNA world to the **non-free-living anaerobic autotrophic LUCA** (Koonin & Martin 2005; Say & Fuchs 2010), the ancestor of prokaryotes, bacteria and archaea (Koonin & Martin 2005). The genome of LUCA was reconstructed by comparing genes present in both archaea and bacteria, which recovered monophyly of Bacterial and Archaeal domains and 355 protein families traced to the very origin of life (Weiss et al. 2016). The origin of bacteria and archaea is still not very well understood. Did the two domains diverge in the very beginning (Weiss et al. 2016), or does one domain stand holophyletic within the paraphyletic parent taxon (Cavalier-Smith 2002)? Biomembranes and cell-walls might provide an answer or deepen the question. **Ether** bounds present in archaeal phospholipids are thought to be very different from bacterial **ester** phospholipid bounds (Koonin & Martin 2005), but bacterial and archaeal **murein cell walls** seem to be of the same origin (Subedi et al. 2021). LECA gave rise to the world of prokaryotes, i.e., the world of **Bacteria** and **Archaea**.

PROKARYOTIC DIVERSITY

The majority of the extant phylogenetic and metabolic diversity belongs to the **prokaryotes** (Figure 2) (Lengeler et al. 1999; Curtis et al. 2002, 2006), the fusion of which, known as symbiogenesis (Martijn & Ettema 2013) or specifically eukaryogenesis (Martin 2017) gave rise to the eukaryotes (Lane & Martin 2015). Prokaryotes are very diverse and inhabit all the places on the planet, from the depths of earth's crust to the top of the atmosphere (Pikuta et al 2007, Burrows et al. 2009). As the topic of this thesis is evolution and reconstruction of the Last Eukaryotic Common Ancestor, here are mentioned just some of the major prokaryotic taxa, some of which were crucial for the origin of the eukaryotes, and the origin of primary plastids respectively (De Alda et al. 2014). Unlike sexual reproduction in eukaryotes, prokaryotes lack anything homologous to sex, but instead, they exhibit **lateral gene transfer** (**LGT**, or **horizontal gene transfer**) (Ku et al. 2015b; Ku & Martin 2016; Garg & Martin 2016; Colnaghi et al. 2020). By this phenomenon, many genes travelled across the prokaryotic tree of life and today those genes represent the major problem of the prokaryotic phylogeny (Martin et al. 2016). LGT is so common in the prokaryotic world that today it is known, for example, that Haloarchaea originated from methanogenic archaea who had a tremendous LGT from many groups of Bacteria (Nelson-Sathi et al. 2012). LGT can be ancient, as in the origin of Haloarchaea, or it can be recent, within the extant taxa (Ochman et al. 2000; Robinson et al. 2013). Extant proteobacteria have approximately seventh of the genome (more specifically 14%) originating from recent LGT event (Kloesges et al. 2011). Test of "verticality" of prokaryotic genes showed that indeed there are certain protein families which are not so prone to LGT, and thus can be used to resolve the prokaryotic phylogeny (Nagies et al. 2020).

Prokaryotes are divided into two large domains, **bacteria** and **archaea** (Woese et al. 1990; Williams et al. 2020). The ancestor of bacteria is known as **LBCA** (Last Bacterial Common Ancestor), while the ancestor of archaea is **LACA** (Last Archaeal Common Ancestor) (Makarova et al. 2007; Xavier et al. 2021). The two domains are separated on the very origin of life and they represent descendants of the first two populations of living cells (Koonin & Martin 2005). Among the well-known groups of bacteria are **proteobacteria** (Spain et al. 2009), which gave rise to the mitochondria (Fan et al. 2019; Muñoz-Gómez et al. 2021), and **cyanobacteria** (Hammerschmidt et al. 2021), which gave rise to the primary plastids (Ponce-Toledo et al. 2017). The mitochondrion is literally a proteobacterium living inside a eukaryotic cytosol, just as a plastid is literally a cyanobacterium that survived inside this very cytosol (Gould et al. 2008). Besides the groups involved in eukaryogenesis, one of the most morphologically complex bacterial assemblages are multicellular *Actinobacteria* and *Myxococcales*, which have a complex lifecycle, superficially similar to the eukaryotic one (Kaiser et al. 2003, van Bergeijk et al. 2020).

The diversity of bacteria is much better understood than the diversity of archaea (DeLong & Pace 2001). However, in recent times, new groups of bacteria are being discovered, but without cultivated members, thus being named "**candidate phyla**" (Hug et al. 2016).

Members of the archaeal domain were considered, until recently, extremophiles (DeLong & Pace 2001), but with novel methods of discovering hidden diversity, such as **metagenomics**, archaeal lineages were discovered everywhere (Munson et al. 1997, Eme & Doolittle 2015; Söllinger & Urich 2019). Besides well-known *Euryarchaeota* which include Haloarchaea and many methanogens, one of the most interesting groups to be recently discovered is certainly ***Asgard*** archaea, regarded to be the closest relatives to the host who gave rise to LECA (Sprang et al. 2015; Zaremba-Niedzwiedzka et al. 2017). Even though the quality of the MAGs (Metagenome Assembled Genomes) from metagenomics is still questionable (Garg et al. 2021), it shows that there is a huge diversity of genes not belonging to any of the known prokaryotic lineages (Hug et al. 2016). Furthermore, besides the metagenomic methods, the first Asgard species were recently cultivated and found to be in obligatory syntrophy with other archaea and bacteria (Imachi et al. 2020). Such a syntrophic lifestyle probably harbours many predispositions that led to eukaryogenesis (Imachi et al. 2020; Schleper & Sousa 2020).

"**Candidatus *Prometheoarchaeum syntrophicum* strain MK-D1**" is, for now, the first and the only Asgard member to be cultivated, and is so peculiar that it took 12 years to get an almost pure culture (Imachi et al. 2020). An interesting fact is that it was sampled from the ocean much before Asgard metagenomic sequences were published. One cell division in this astonishing organism takes two weeks, so the growth of the culture is very slow (Imachi et al. 2020). The key to the understanding of eukaryogenesis certainly lies among such deep-sea organisms, only few of which are known to be cultivated at the moment of writing this thesis (Zhang et al. 2018; Imachi et al. 2020).

Syntrophy is not rare in the prokaryotic world (López et al. 2010). There are many examples besides the "Candidatus *Prometheoarchaeum syntrophicum"*. Many prokaryotes live as **colonies**, **mats**, or **biofilms** (Costerton et al. 1995; López et al. 2010). Many mats and biofilms have a complex structure (Stolz 2000; Futo et al. 2021). Among the diverse syntrophic relationships within and between bacteria and archaea, two billion years ago (Porter 2020) one gave rise to the endosymbiotic event which resulted in the origin of the third domain of life, the eukaryotes (Koonin & Martin 2005, Zimorski et al. 2014).

ORIGIN OF EUKARYOTES

Eukaryotes were once regarded a separate domain equal in rank to Bacteria and Archaea (Woese et al. 1990), however it is an accepted view in modern evolutionary biology that they represent only a **hybrid domain** that originated by polyphyletic event from two distinct prokaryotic parents (Brückner & Martin 2020), an **archaeal host** and a **proteobacterial endosymbiont** (Roger et al. 2017; Fan et al. 2020; Zimorski et al. 2014; Imachi et al. 2020). The origin of the eukaryotes, or eukaryogenesis, marks the transition between the First Eukaryotic Common Ancestor (**FECA**) and the Last Eukaryotic Common Ancestor (**LECA**) (Butterfield 2015; O'Malley et al. 2019). Many hypotheses have hitherto tried to explain the FECA transition to LECA, but also many have failed (e.g., Cavalier-Smith 1987; Margulis et al. 2000). It is clear today that one needs to seek the eukaryotic parents among the (extant) prokaryotic lineages. As the eukaryotic genome indeed represents a chimera of archaeal and bacterial genes (Brückner & Martin 2020), homologues of the prokaryotic-derived genes are being defined all the time among the eukaryotes (Timmis et al. 2004; Zaremba-Niedzwiedzka et al. 2017; MacLeod et al. 2019; Imachi et al. 2020).

Some hypotheses on the origin of eukaryotes did not take into account the symbiogenic nature of this event (Gould & Dring 1979; Fournier & Poole 2018), and those are hence rejected today as they failed to explain the chimeric nature of the genome. Not only is hybrid genome evidence of the symbiogenesis, but there is also the mitochondrion, a true proteobacteria-descendant-organelle. It is basically an autonomous mutualistic bacterium living in the eukaryotic cytosol (Henze & Martin 2003; Roger et al. 2017). Only endosymbiosis provides a plausible explanation of how one prokaryotic cell, nowadays called an organelle, could have entered another prokaryotic cell (archaeal host) and survived inside of it (Martin & Müller 1998; Zimorski et al. 2014; Imachi et al. 2020). Some of the famous **mutualistic hypotheses** (Figure 3), including two (see **1, 2, 3, 6**) or even three prokaryotic lineages (see **2, 4**, **5** and, **7**), are, by name:

(1) the hydrogen hypothesis explains how LECA came to be through a biochemical symbiosis, i.e., anaerobic syntrophy, involving an autotrophic anaerobic hydrogen-dependent archaebacterial host and a bacterial endosymbiont who generated hydrogen as a waste. This bacterial endosymbiont gave rise to mitochondria, as well as to hydrogenosomes (first column in Figure 3A) (Martin & Müller 1998, Martin 1999)**.**

(2) the 'classical' endosymbiotic hypothesis (Margulis et al. 2006; De Duve 2007), **the symbiotic association hypothesis** (Vellai & Vida 1999), and the **ox-tox hypothesis** put the emphasis on oxygen uptake and state that an anaerobic archaeon engulfed an endo-symbiont with respiration ability, who became mitochondrion (second column, Figure 3A).

(3) the sulphur cycling hypothesis explains how the origin of the eukaryotes likely started with mutualism of H_2S-producing archaeal host and H_2S-consuming alphaproteobacterial endosymbiont, which became an organelle, i.e., the mitochondrion (third column in Figure 3A) (Searcy 1992).

(4) the syntrophy hypothesis interprets the eukaryogenesis as an event in which a methanogenic euryarchaeon entered a fermentative myxobacterium, and this was followed by another endosymbiosis, the one in which a protomitochondrion (i.e., alphaproteobacteria) entered this chimeric cell and by compartmentalization gave rise to LECA (fourth row in Figure 3A) (López-García & Moreira 2006).

(5) the phagocytosing archaeon hypothesis states that an archaeal host engulfed, by phagocytosis, endosymbiotic bacteria, and this event resulted in the formation of the nucleus. Furthermore, protomitochondrion entered this nucleate cell and survived inside of it as a mitochondrion (fourth row in Figure 3A) (Martijn & Ettema 2013). Hypotheses (4) and (5) are known as mitochondria-late hypotheses, as they explain that mitochondrion was not a necessary predisposition for the formation of LECA.

(6) the inside-out hypothesis explains the origin of complex eukaryotic structure through the mutualism of the bacterial symbionts living on the cell surface of the archaeal host, which later formed a net of membranes around the epi-symbionts and consequently the outer membrane originated, giving rise to LECA or the first eukaryotic cell (fourth row in Figure 3A) (Baum & Baum 2014).

(7) the **E3 hypothesis or the 'Entangle-Engulf-Endogenise model'** is the most recent eukaryo-genesis model, which explains how LECA originated from an Asgard host similar to the 'Candidatus *Prometheoarchaeum syntrophicum*' *via* syntrophy with sulphate-reducing, and aerobic organotrophic bacteria, respectively (Figure 3B) (Imachi et al. 2020).

With so many, sometimes contrary, assumptions and hypotheses being published continuously during the last decades, it is not easy to get a clear picture on what is fact about the eukaryotic ancestor, and what are only hypotheses to be tested. We now present some of the strongly supported observations and inferences about the first eukaryote. The Last Eukaryotic Common Ancestor (LECA) was a nucleate, mitochondriate and sexual heterotrophic organism that originated via eukaryogenesis, from mutualistic relationship of phylogenetically distinct prokaryotes (archaeal host and proteobacterial endosymbiont) all to the self-sufficient cell with a complex eukaryotic lifecycle that involves mitosis, and meiosis i.e., reciprocal recombination which is the preadaptation for sex. LECA had complex endomembrane system, composed of nucleus, Golgi apparatus, and smooth and rough endoplasmic reticulum (e.g., Speijer et al. 2015; Garg et al. 2016; O'Malley et al. 2019; Hofstatter & Lahr 2019). Such a complex endomembrane system likely originated from vesicular activity of the early proteobacterium that entered an archaeal cytoplasm (Gould et al. 2016).

EUKARYOTIC DIVERSITY

Eukaryotic diversity is enormous (Adl et al. 2019) and with several million described species (CoL 2022), this is phenotypically the most diverse domain of life (Figure 4., Figure 5.) (Archibald et al. 2017). However, the majority of the eukaryotic lineages are most probably not yet discovered. The deep ocean hides many obscure groups (Cordier et al. 2022), of which some are today regarded in the protistology as "small" or "not-well-known" (Adl et al. 2012, 2019), but future research could find out that those "small" groups represent phylogenetically the most diverse ones on our planet (Cordier et al. 2022). Well-known eukaryotic groups are, of course, those that people often see and those with which people often work and this is *why* in the beginning of natural classification all the living beings were divided into two macroscopic groups – non motile plants (**Vegetabilia**) and motile animals (**Animalia**) (Linné 1735). Fungi were already in the 18th century, in comparison with animals and other plants, regarded obscure, so Linné (1751), in the very beginning of their classification already wrote *'Fungorum ordo (...) Chaos est (...)'* or in free translation *'The order of fungi is chaotic'* (Kowallik & Martin 2021). As more groups were discovered with developments in the methods of fieldwork, material preservation, microscopy and molecular biology, more and more new groups were added to the natural system (Archibald et al. 2017; Adl et al. 2019). **Protists** were defined in 19th century in order to incorporate all the newly discovered microscopic organism that did not fit into animals, plants or fungi, but were "more primitive" in traits (Hogg 1860; Haeckel 1866).

The most famous and perhaps the most widely used system of the eukaryotic classification did not change a lot since Hogg's (1860) and Haeckel's (1866) system and includes four large **kingdoms**, (i) **animals**, (ii) **plants**, (iii) **fungi** and (iv) **protoctists** (protists) (e.g., Whittaker 1969). It is clear today that three of these groups (animals, plants, and fungi) have their origin within the fourth group, i.e., within the protists (Cavalier-Smith 1981). Specifically, animals are known to be sisters to Choanoflagellates, of which some were regarded "primitive" animals in the past, while others assigned to the protists (e.g., Kent 1880). Plants are in modern taxonomy are understood as a group of land-specialized green algae (Gould et al. 2008; de Clerck et al. 2012; Leliaert et al. 2012; de Vries & Gould 2018). Fungi did not become *less* chaotic during this time, but instead, they have become *trickier* to classify (Hibbett et al. 2007; Spatafora et al. 2016; Adl et al. 2019). Fungi are known to be close to cristidisoideans, which is another obscure group of fungi-like eukaryotes (Adl et al. 2019). The four kingdoms classification was based on simple shared characters (sessile vs. motile; heterotrophic vs. photosynthetic; amoeboid vs. of definite shape etc.) or on the very lack of them and as such is not useful anymore (Cavalier-Smith 1981; Adl et al. 2019). The problem with eukaryotic classification is conflicting data, molecular and morphological. Figures from 4. to 17. and the following text is to serve as a brief atlas to eukaryotic diversity.

In modern systems, eukaryotes are divided into many supergroups and groups (Adl et al. 2012; 2019). All the evolutionary groups are designed to be monophyletic, and the system tends to define more holophyletic and less paraphyletic groups (Adl et al. 2019; Nagies et al. 2020). The modern system was based on a combination of phylogenomic data and comprehensive systemic research on biochemistry, physiology, morphology and ecology, i.e., it represents a holistic approach (e.g., Cavalier-Smith 1981; Archibad et al. 2017; Adl et al. 2019). Many **supergroups** have been defined within the eukaryotic domain, none of which correspond to the old-kingdom-system (Adl et al. 2019; Colp & Archibald 2019). The large and the best-known supergroups are **Amoebozoa** (amobae and mycetozoans), **Archaeplastida** (red algae, green algae, and glaucophytes), **Cryptista** (cryptophytes and relatives), **Discoba** (one part of the excavates), **Haptista** (haptophytes and relatives), **Metamonada** (the other part of the excavates), **Obazoa** (animals, fungi, and relatives), and **SAR** (a diverse group composed of **S**tramenopiles or heterokont algae: such as diatoms and brown algae, but also oomycetes; **A**lveolata: dinoflagellates, apicomplexans, ciliates; and **R**hizaria: foraminifera, radiolarians and their amoeboid relatives) (Archibald et al. 2017; Adl et al. 2019). There are also many, relatively recently discovered groups, such as Hemimastigophora (Lax et al. 2018), Collodictyonida (Zhao et al. 2012), or Telonemida (Yabuki et al. 2013; Strassert et al. 2019) (Figures 4, 5), that did not fit into those well-known supergroups, so the definition of certain taxa has been amended (compare Adl et al. (2012) with Adl et al. (2019)). Obazoa and SAR are good example of the dynamics of the system of eukaryotic classification. Obazoa is basically a taxon established to include Opisthokonta and their two not-well-known sister groups, Breviatea and Apusomonadida (Adl et al. 2019). Obazoa might be regarded as 'wider Opisthokonta' (Adl et al. 2012, 2019). The same is true with TSAR, a newly defined clade to include Telonemida and SAR (Figure 4) (Tikhonenkov 2020).

The **root** of the eukaryotic tree has not yet been clearly defined (Simpson & Roger 2002; Brinkman & Philippe 2007). Some analyses place Amoebozoa and Opisthokonts together, as the eukaryotic root and because of that, all the eukaryotes can be divided into two large groups. Basal **Amorphaea** comprises Amoebozoa and Obazoa, while crown **Diaphoretickes** includes Archaeplastida, Cryptista, Haptista, Excavata (Discoba and Metamonada), and TSAR (see Figure 4 for the full system of modern eukaryotic classification) (Adl et al. 2012). Some authors have suggested that amitochondriate Archaezoa might be the first branching clade within the eukaryotes (Cavalier-Smith 1989), but now it is known that the lack of mitochondria represents an apomorphic, and not plesiomorphic trait within the eukaryotic domains (Martin & Müller 2007). Some eukaryotes secondarily lost the function of mitochondria and have mitosomes or have completely lost these organelles (Roger et al. 2017). The root of the eukaryotes might also lie within Excavata, or this whole clade might be the root (He et al. 2014). Because of the taxon-rich sampling performed in this thesis, the most interesting candidate for the eukaryotic root is Opishokonta, which came out as the root in other taxon-rich analyses (Stechmann & Cavalier-Smith 2002; Cerón-Romero et al. 2021). In order to be able to answer questions about eukaryotic origin and LECA, it is important to understand modern eukaryotic diversity to its full extent, because only taxon-rich sampling in such a diverse group can result in systematic conclusions (Parfrey et al. 2010; Katz & Grant 2015).

The diversity of the supergroup **Amoebozoa** (Figure 6) is not yet well understood. Many representatives of this supergroup live in deep sea habitats (Kudryavtsev & Pawlowski 2013; Volkova & Kudryavtsev 2017), but besides the parasitic forms (e.g., Clark et al. 2006) and famous multicellularity models, such as *Dictyostelium discoideum* (e.g., Loomis 2012), not a lot of information is available on those interesting, ecologically abundant, phagotrophic, and usually amorphic eukaryotes (Cavalier-Smith et al. 2004; Archibald et al. 2017; Adl et al. 2019). The cytosol of many species harbours many nuclei or they exhibit some kind of a multinucleate (syncytial) phase (Adl et al. 2019). The historical group of 'protist amoebae' does not correspond to modern Amoebozoa. Many amoebozoans are sexual, but certain lineages lost sexual reproduction in the distant past. They have escaped accumulation of the lethal mutations caused by the lack of recombination, a phenomenon known as the Muller's (1932) ratchet, because of a high polyploidy level, which can delay the effects of Muller's ratchet (Kondrashov 1994). Some species might have the whole genome duplicated several dozen thousand times (Maciver 2016). Amoeboid morphology is, just like mycoid morphology, widespread in the eukaryotic tree and "old amoebae" have been scattered all over the eukaryotic tree now in the last few decades (Adl et al. 2012, 2019; Archibald et al. 2017).

The supergroup of **Obazoa** (opisthokonts, plus apusomonadids and breviates) (Figure 7, Figure 8) are far more studied members of the amorphaean clade than Amoebozoa. **Opisthokonts** include heterotrophic osmotrophic **Fungi** (Figure 8) and **Animals** (Figure 9, Figure 10), for which the previous sentence is especially true. The definition of animals is simple. They are eukaryotes which undergo morula and blastula as two very characteristic phases of embryonic development (Nielsen 2012). Colonial choanoflagellates are a sister group to animals (Figure 7), but it is not clear yet whether the colonial structure has been inherited from the common ancestor, or is a derived structure. Besides Fungi and Animalia, the supergroup Opisthokonta includes some "smaller phyla", so it is usually divided into two subgroups – **Holozoa** or animals and relatives; and **Holomycota**/Nucletmycea or fungi and relatives. If opisthokonts really represent the eukaryotic root, or if the eukaryotic root lays within the opisthokonts, it is clear that much of the ancestral eukaryotic morphology and physiology, i.e., LECA's traits, need to be searched for among the diversity of extant representatives of this group (Stechmann & Cavalier-Smith 2002; Cerón-Romero et al. 2021). Most fungi exhibit multinucleated hyphae, most animals have some multinucleated tissues, and many fungi and animals' relatives exhibit multinucleate phases as well (Kiss et al. 2019; Adl et al. 2019). Animals are the most species-rich taxon on earth, with more than a million described species, but some estimate that several million fungal species may be still undiscovered, so it remains a question which group on Earth is currently the most diverse (Wu et al. 2019). If the whole problem of the **definition of species** (Wilkins 2018) is added to the equation, it is clear that it will never be known how many species could (have) be(en) on the planet, but also it is not even important to answer, as there is almost no meaning in the term *species* if all the eukaryotes are taken into account. The concept of species, especially the biological one, is almost useless in the microbial world, not to mention the prokaryotic world. It was made, almost exclusively, for animals (Mayr 1942). However, without species names we cannot communicate, so even the concepts may be problematic, the principle is essential.

The **supergroup** of **Excavata** (Figure 11) is likely not monophyletic, but composed of two very distant clades, **Discoba** and **Metamonada** (Adl et al. 2012, 2019). Relationships between and within those two groups are not yet resolved (Kamikawa et al. 2014; Simpson et al. 2017; Archibald et al. 2017). Besides the **photosynthetic** euglenoids and **free-living** forms, such as *Dysnectes*, many groups are obligately **parasitic** (see the diversity of the group in Figure 11). Famous examples of parasitic members are genera *Trypanosoma,* member of Kinetoplastea (El-Sayed et al. 2005; Borges et al. 2021), and *Giardia*, member of Diplomonadida (Plutzer et al. 2019). Most excavates are unicellular, but there are multicellular examples, such as Acrasidae (Brown et al. 2012; Adl et al. 2019). Despite their unicellular nature, most members are multinucleate, and their usually harbour two or four nuclei in the cytoplasm (Archibald et al. 2017; Adl et al 2012, 2019). Furthermore, the astonishing variation within this group that is sometimes regarded basal within the eukaryotes (He et al. 2014) is evident in the structure of mitochondria. Members of Excavata vary from completely reduced mitosomes, i.e., mitochondria that are so reduced that they lack even DNA, to the Jakobids which have the largest mitochondrial genomes among the Eukarya (Strassert et al. 2016).

The **supergroup** of **Archaeplastida** gathers eukaryotes whose ancestor was involved in the endosymbiotic event in which a population of cyanobacteria survived inside of its cytoplasm and became plastids (Figure 12) (Gould et al. 2008). The primary plastid is the one with two membranes. The glaucophyte plastid is interesting, as it is the only plastid in the eukaryotic domain that still has a peptidoglycan layer between the membranes. It is not clear yet whether this plastid represents a plesiomorphous, ancestral type, or a very derived one (Figueroa-Martinez et al. 2019). All members of the Archaeplastida lineage are autotrophic photosynthetic eukaryotes with plastids, except for the recently discovered Rhodelphidia, sister to red algae, which are phagotrophic and lack plastids (Gawryluk et al 2019). Archaeplastida includes (Figure 13, Figure 14) two large groups, red algae (**Rhodophyta**) and green algae (**Chloroplastida**), and one small group, glaucophytes (**Glaucophyta**). Relationships between those groups, as is the case with many eukaryotic groups, are not clear (De Clerck et al. 2012, de Vries & Gould 2018). Many known archaeplastid species are multicellular or multinucleated, but there are also many unicellular forms present in the sea and in the freshwaters (Archibald et al. 2018). Plants, that were until recently regarded a separate kingdom, are today known to be specialized green algae that inhabited the land. Plants are related to *Zygnematophyceae*, *Charophyceae*, and their other freshwater relatives together with whom they form a group of *Streptophyta* (Figure 12).

Unlike the endosymbiotic events with green algae that involved the ancestors of Euglenophyceae (Excavata: Discoba) and *Chlorarachniophyceae* (SAR: Rhizaria), ancestor of the **supergroup Hacrobia** (Figure 15), i.e., that of **Haptista** and **Cryptista**, was involved in the endosymbiotic event with a population of red algae (Gould et al. 2008). Certain species, for example, of coccolithophore exhibit changes between haploid and diploid phase in their lifecycles. The diploid phase, often covered with calcified scales is known to consume much more energy than the haploid phase, which is often amoeboid. Cryptomonads are known to have a nucleomorph, a remnant of the endosymbiont's nucleus, inside of the complex plastid,

residing in the periplasmatic space between the two outer plastid membranes and the two inner plastid membranes. (Eikrem et al. 2017; Hoef-Emden & Archibald 2017).

The supergroup of **SAR** (abbreviation of the first letters of the three groups included – **Stramenopiles**, **Alveolata**, and **Rhizaria**) (Figure 16, Figure 17). This supergroup is, by the variety of its members, the closest one in the modern eukaryotic classification system to what was Protista in the old system (Haeckel 1866; Adl et al 2012, 2019). SAR gathers many amoeboid members, for example vampyrellids and foraminifers; many mycoid members, for example oomycetes and labyrinthulids; as well as many plant-, i.e., algae-like eukaryotes (Figure 17). Photosynthetic members of SAR are numerous heterokonts (Ochrophyta), but also some dinoflagellates and ciliates. The heterokont ancestor was a heterotrophic eukaryote who engulfed red algae. This ancestor gave rise to diatoms, brown algae and many related lineages. A rhizarian engulfed a population of green algae and became the ancestor of chlorarachniophyceans, the only photosynthetic rhizarians with secondary plastid (Gould et al 2008). There is one more photosynthetic rhizarian, but this one is with primary plastid, independent from one in Archaeplastida – amoeboid *Paulinella* (Lhee et al. 2019). The supergroup of SAR if full of bizarre examples not seen anywhere else in the eukaryotic world. The first example is *Pseudoblepharisma tenue,* a ciliate (Alveolata: Ciliata) with two autotrophic endosymbionts. One is a prokaryotic photosynthetic bacterium (not a cyanobacterium, but a candidate gammaproteobacterium), while the other is a green alga (Muñoz-Gómez et al. 2021). Many ciliates are known to have a generative micronucleus devoted to heredity and a vegetative macronucleus devoted to gene expression, with a macronuclear genome composed of around 30 million nanochomosomes, or more specifically 16 000 nanochromosomes in the haploid set, with the polyploidy level of 1 900 times (Swart et al. 2013; Kumar & Kumari 2015). Eukaryotes are very diverse, and it is not an easy task to find ancestral traits among the modern diversity.

Others would say they are the strongest monophylum in nature, but they are not really monophyletic. All eukaryotes have nuclei, as well as complex inner endomembrane system including Golgi apparatus and endoplasmic reticulum, all made of bacterial phospholipids (Embley & Martin 2006; Gould et al. 2016). For many unique structures, genes and processes that occur only in this domain, but not in all of its members, it was found that they represent ancestral traits likely present in LECA. Most eukaryotes are, like their ancestor, heterotrophic (Adl et al. 2019), but again, there are those who have photosynthesizing cyanobacteria living in their cytosol and functioning as an organelle, plastid (Gould et al. 2008). Those autotrophic eukaryotes are known to be derived, specialized forms. Most eukaryotes are sexual, but there are many examples of asexual forms. An interesting example is polyploid asexual amoebae, which escaped Muller's ratchet without recombination, but with many thousands of genome copies (Maciver 2016). Eukaryotes are ancestrally sexual. Full meiotic machinery is ancestral, so asexual eukaryotes are definitely derived, specialized lineages (Hofstatter & Lahr 2019). Most eukaryotes have mitochondria, but some do not (Roger et al. 2017). Those who lack mitochondria are proven to have lost them, evidencing that the eukaryotic ancestor was indeed mitochondriate cell (Martin & Müller 2007). This thesis questions, among other things, whether eukaryotes are ancestrally multinucleated or as usually depicted, uninucleate.

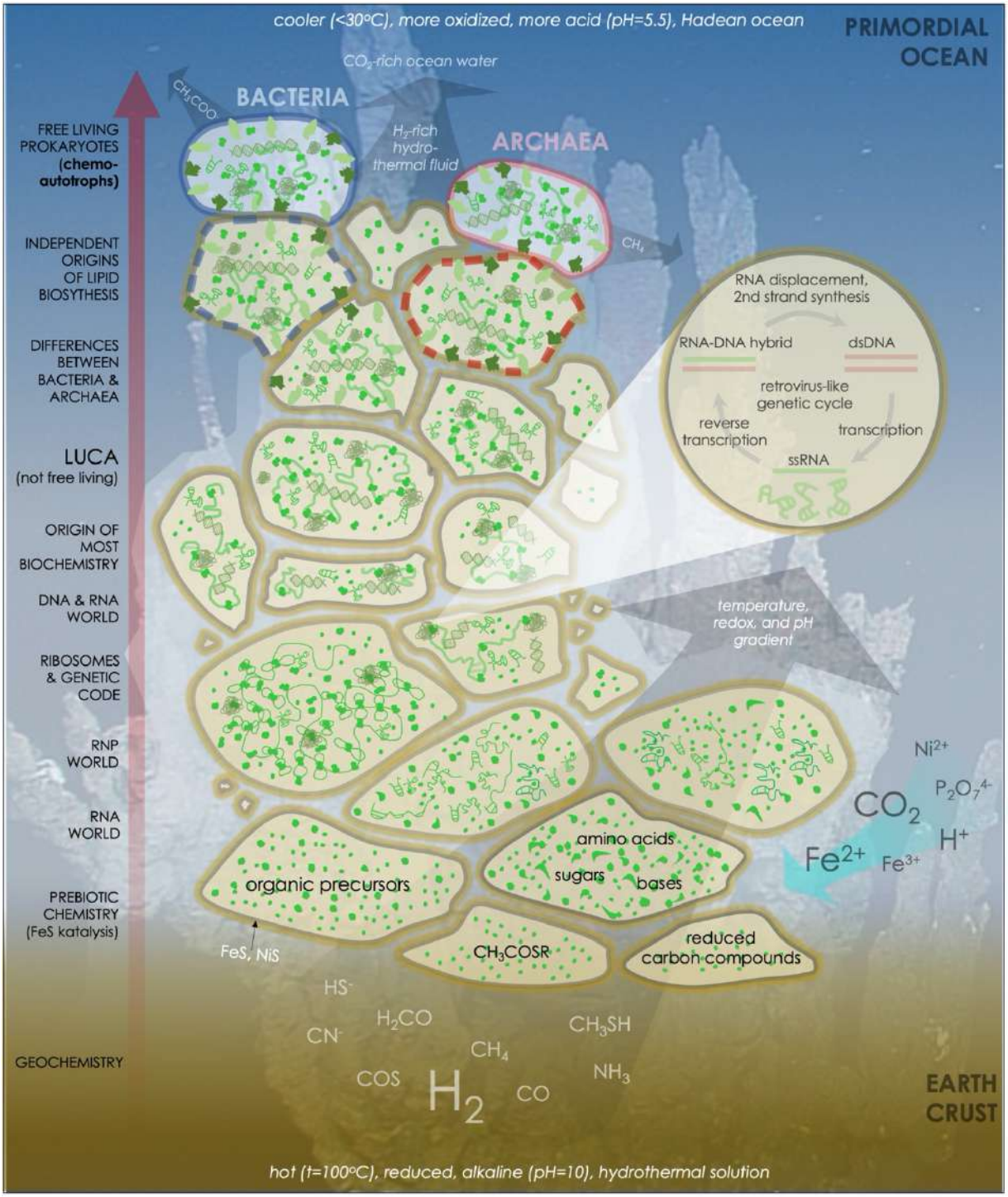


Figure 1. Schematic visualization of the origin of life. The timeline shows the transition between geochemistry and the first prokaryotes, at a hydrothermal vent at the bottom of the ocean, some 3.8 billion years ago. Drawn after Martin & Russell (2003) and Koonin & Martin (2005). Hydrothermal current (hot, alkaline solution rich in hydrogen) penetrates rocks built out mostly of pyrite. Sea current coming from the primordial ocean is on the OTHER HAND rich in carbon dioxide, cooler, and more oxidized. There, on the contact of the two, temperature, redox, and pH gradients occur, so from geochemistry originated the biochemistry, present in all the living cells. Source of the photographs of the hydrothermal vent in the background Wikimedia commons, NOAA Okeanos Explorer Program, Galapagos Rift Expedition 2011, LICENCE CC-BY-SA 2.0.

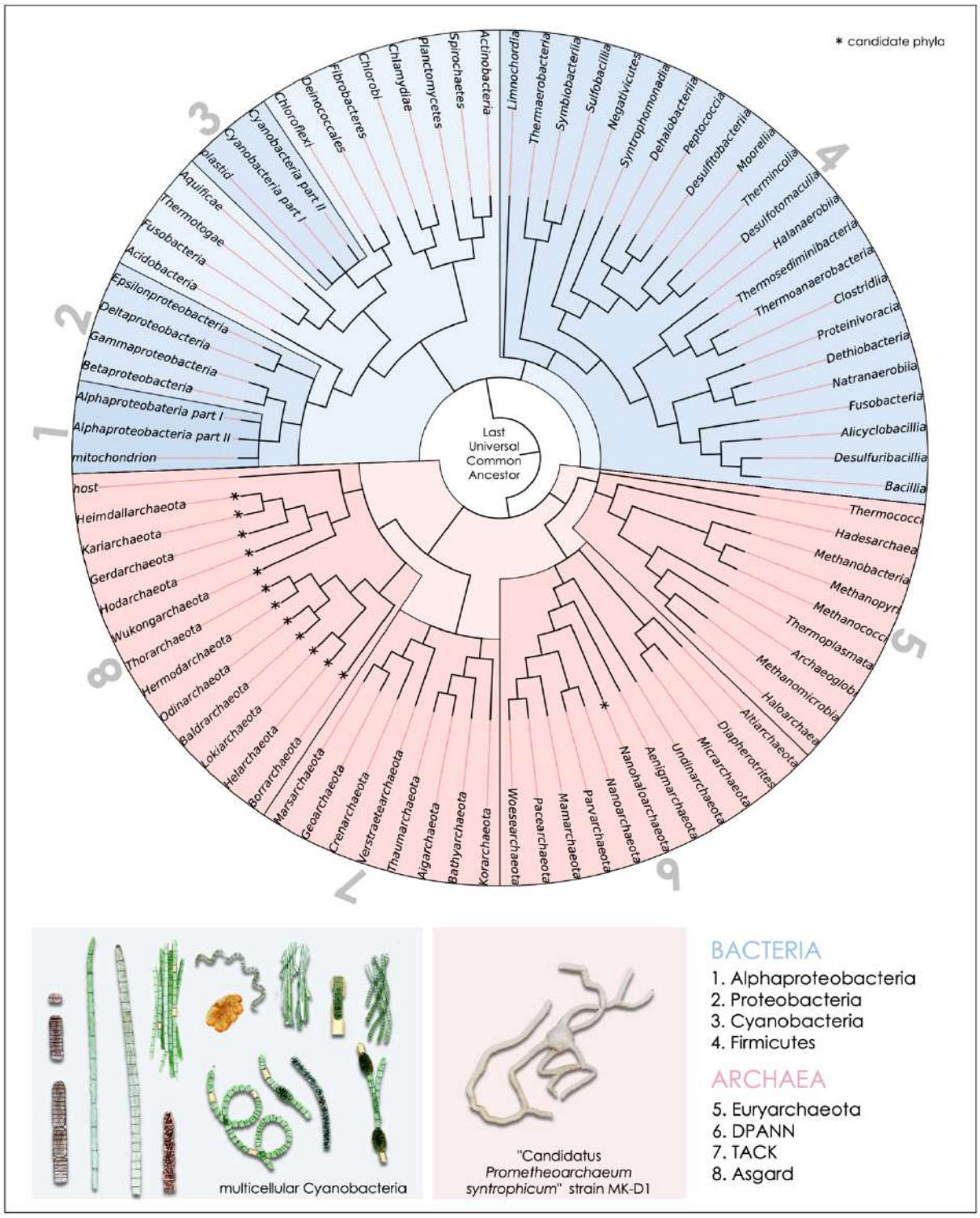


Figure 2. Phylogenetic diversity of prokaryotes, i.e., members of the domains Archaea and Bacteria. Diversity of the Bacteria follows Hug er al. (2016), Mendler et al. (2019), Xavier et al. (2021). For Firmicutes, some obsolete phylo are shown in order to point out the unknown majority, present also in other bacterial phyla. Diversity of the main groups of Archaea follows Baum & Baum (2020), Castelle et al. (2015), Cavalier-Smith et al. (2020), Dombrowski et al. (2020), Eme et al. (2017), Imachi et al. (2020), Bird et al. (2017), Liu et al. (2021), McKay et al. (2019), Williams et al. (2017), Williams et al. (2020). Asgard clades represent candidate phyla, i.e., taxa known from metagenomic assemblages only, without cultivated representatives. Sole example of the cultivated Asgard archaeon is "Candidatus Prometheoarchaeum syntrophicum" strain MK-D1 (source Wikimedia commons, author Maulucioni, licence CC BY-SA 4.0). Bacterial examples depicted are simple filamentous Cyanobacteria (genera *Anabaena*, *Aphanizomenon*, *Nostoc*, *Cylindrospermum*, *Lyngbya*, *Microcoleus*, *Oscillatoria*, *Trichodesmium*, *Phormidium*, and *Spirulina*) (source Wikimedia Commons, Allan Pentecost (2016) Freshwater Biological Association [publ.], environmentdata.org/archive/fbaia:2442, licence CC BY-SA 3.0).

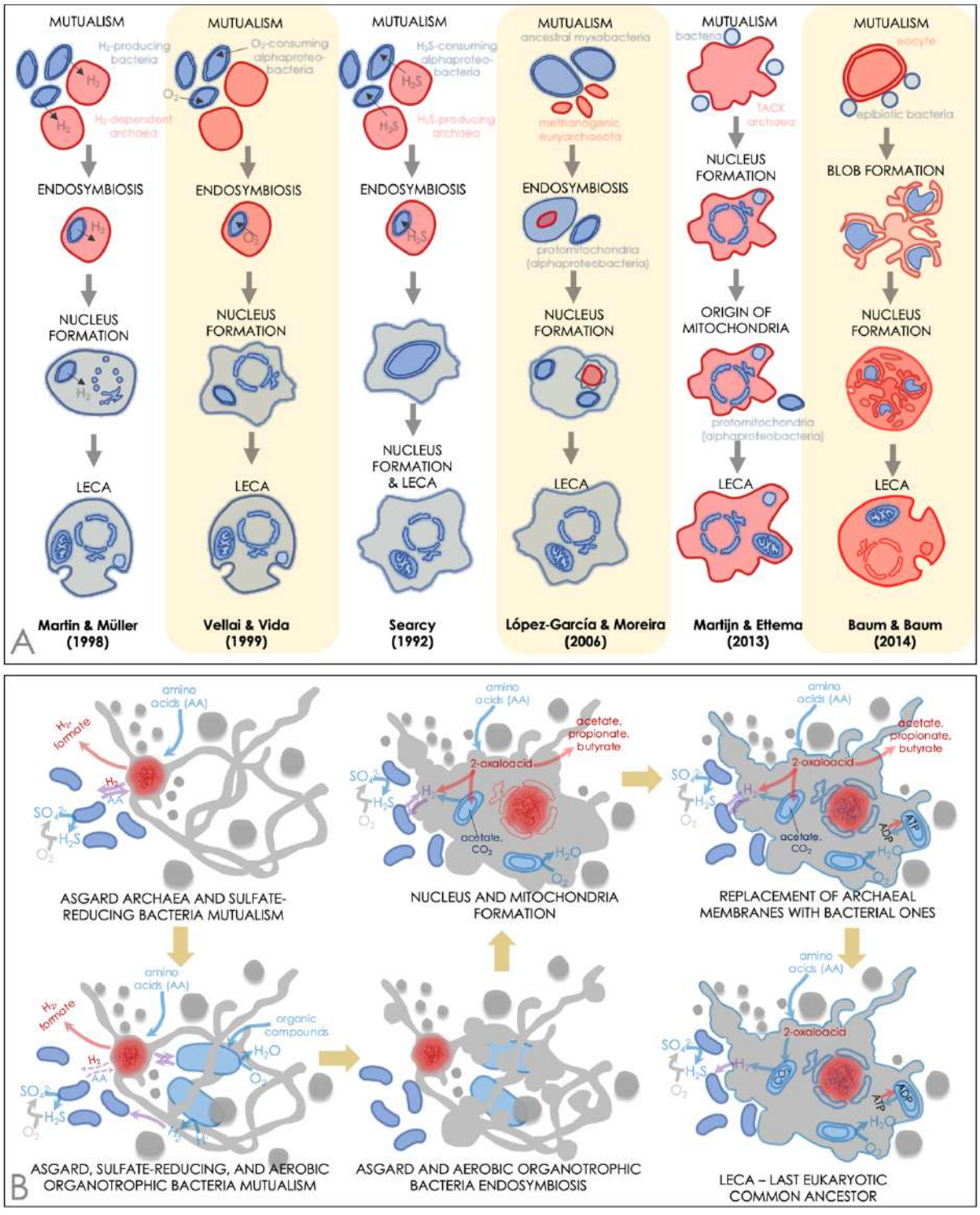


Figure 3. Overview of the mutualistic hypotheses on the origin of eukaryotes, i.e., LECA – the Last Eukaryotic Common Ancestor. Hypotheses that do not include mutualism/syntrophy, such as Raff & Mahler (1972), Bogorad (1975), or Gray (2014), are excluded. **A.** Some of the famous mutualistic hypotheses for the origin of LECA, such **as the hydrogen hypothesis** (Martin & Müller 1998, Martin 1999)**, the symbiotic association hypothesis** (Vellai & Vida 1999), **the sulphur cycling hypothesis** (Searcy 1992), **the syntrophy hypothesis** (López-García & Moreira 2006), **phagocytosing archaeon hypothesis** (Martijn & Ettema 2013), and **the inside-out hypothesis** (Baum & Baum 2014). Drawn by the author after Imachi et al (2020) and Martin et al. (2015). **B. E3 hypothesis or the "Entangle-Engulf-Endogenise model"** hypothesis on the eukaryotic origin (from Imachi et al. 2020) explain how LECA originated from Asgard, Candidatus *Prometheoarchaeum syntrophicum*' strain MK-D1-like ancestor *via* syntrophy with sulphate-reducing, and aerobic organotrophic bacteria, respectively.

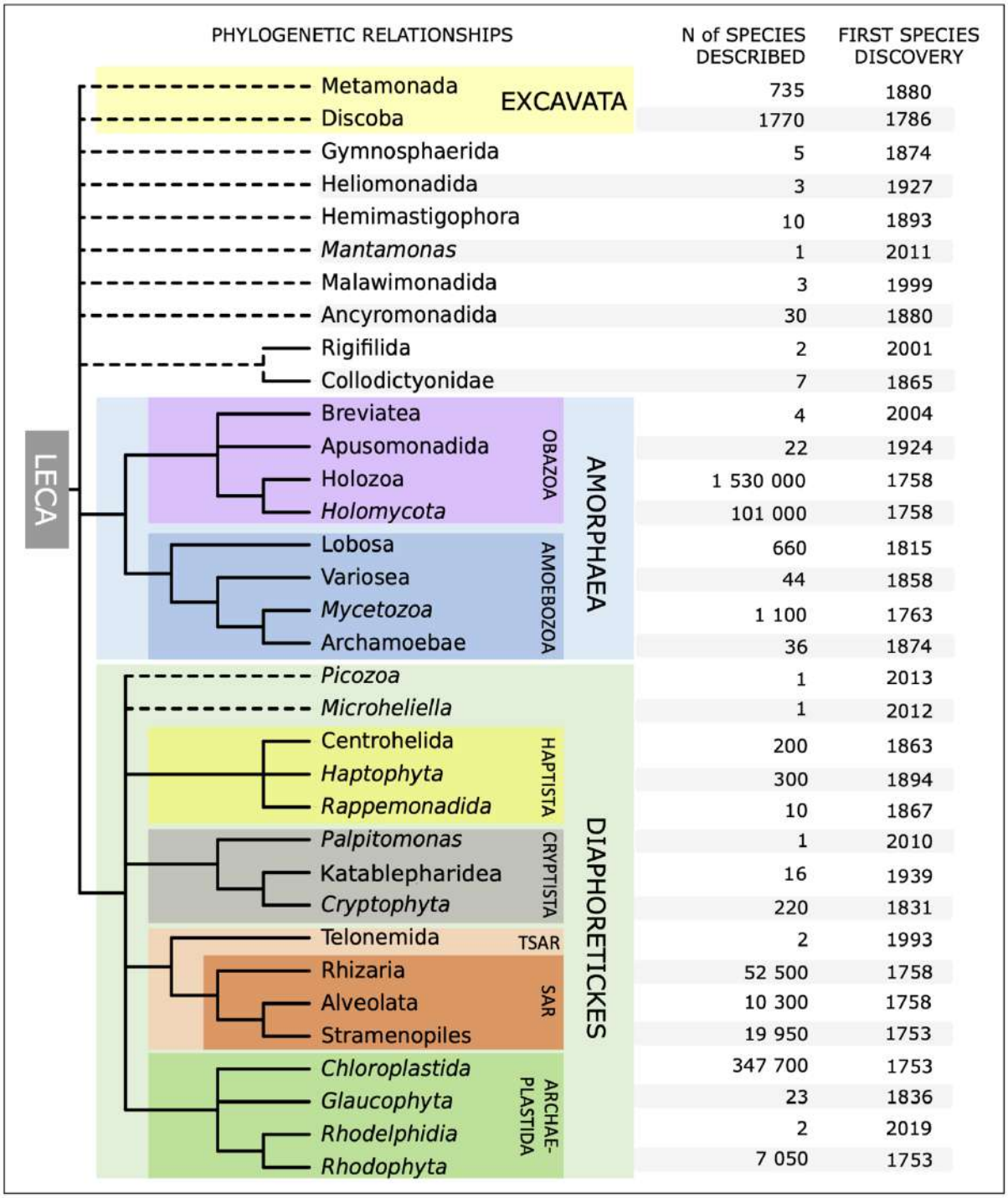


Figure 4. A glimpse into phylogenetic diversity of Eukarya. Amorphaea and Diaphoretickes are two major groups within the Eukaryotes. Full lines represent taxa with more or less clear relationships towards each other's. With dotted lines shown are the groups whose position in the tree of eukaryotes has not been consistent yet (for example Malawimonadida, Hemimastigophora, or Heliomonadida). Together with each taxon, approximate number of the species described is shown, together with the year of the first species' description. Taxa that were described later have much less defined species in comparison with the taxa whose first representatives were described several centuries ago. Holozoa represent Animals and their closest relatives. Among animals, insects have the largest number of described species, currenly more than million. Relationships shown on the cladogram follow Adl et al. (2012, 2019), Berney et al. (2015), Brown et al. (2013), Brueckner & Martin (2020), Burki et al. (2016), Cavalier-Smith et al. (2014), Colp & Archibald (2019), Kamikawa et al. (2014), Leger et al (2017), Leliaert et al. (2012), Simpson et al. (2017), Wickett et al. (2014), and Yabuki et al. (2014).

Figure 5. A glimpse into the morphological diversity of the 'not-well-known' or 'small' eukaryotic taxa. Small phyla are comprised mostly of unicellular and uninucleate cells. A. *Heliomorpha mutans* (source EoL, author Proyecto Agua, licence CC-BY-NC SA); **B. Hemimastigophora gen. sp.** (source EoL, authors A. Feng, Wie-Song and D. J. Patterson, licence CC-BY-NC); **C. *Malawimonas jakobiformis*** (source Wikimedia commons, authors D. J. Patterson, L Amaral-Zettler, M. Peglar and T. Nerad, licence CC BY-SA 3.0); **D. *Picomonas judraskeda* (**source Wikimedia commons, authors Ramkumar Seenivasan, Nicole Sausen, Linda K. Medlin, Michael Melkonian, CC-BY 2.5 Generic); **E. *Collodictyon* sp**. (source EoL, author micro*scope, licence CC-BY-NC); **F. *Ancyromonas* sp**. (source EoL, authors Kathy Sheehan and David Patterson, licence CC-BY-NC); **G. *Raphidiophrys contractilis*** (source Wikimedia common, author ja:User:NEON / commons:User:NEON_ja - Own work, CC BY-SA 2.5); **H. Amastigomonas sp.** (source EoL, author Naja Voers, licence CC-BY-NC); **I. *Telonema* sp.** (author Mats Kuylenstierna, copyright holder Maria KuylenstiernaCC-BY SA 3.0); **J. *Breviata anathema* (**source EoL, authors David Patterson, Linda Amaral Zettler, Mike Peglar and Tom Nerad, licence CC-BY NC); **K. *Roombia truncata*** (source EoL, authors Okamoto N, Chantangsi C, Horák A, Leander B, Keeling P, licence CC-BY 3.0); **L. *Acanthocystis turfacea*** (source EoL, author Proyecto Agua, licence CC BY-SA NC).

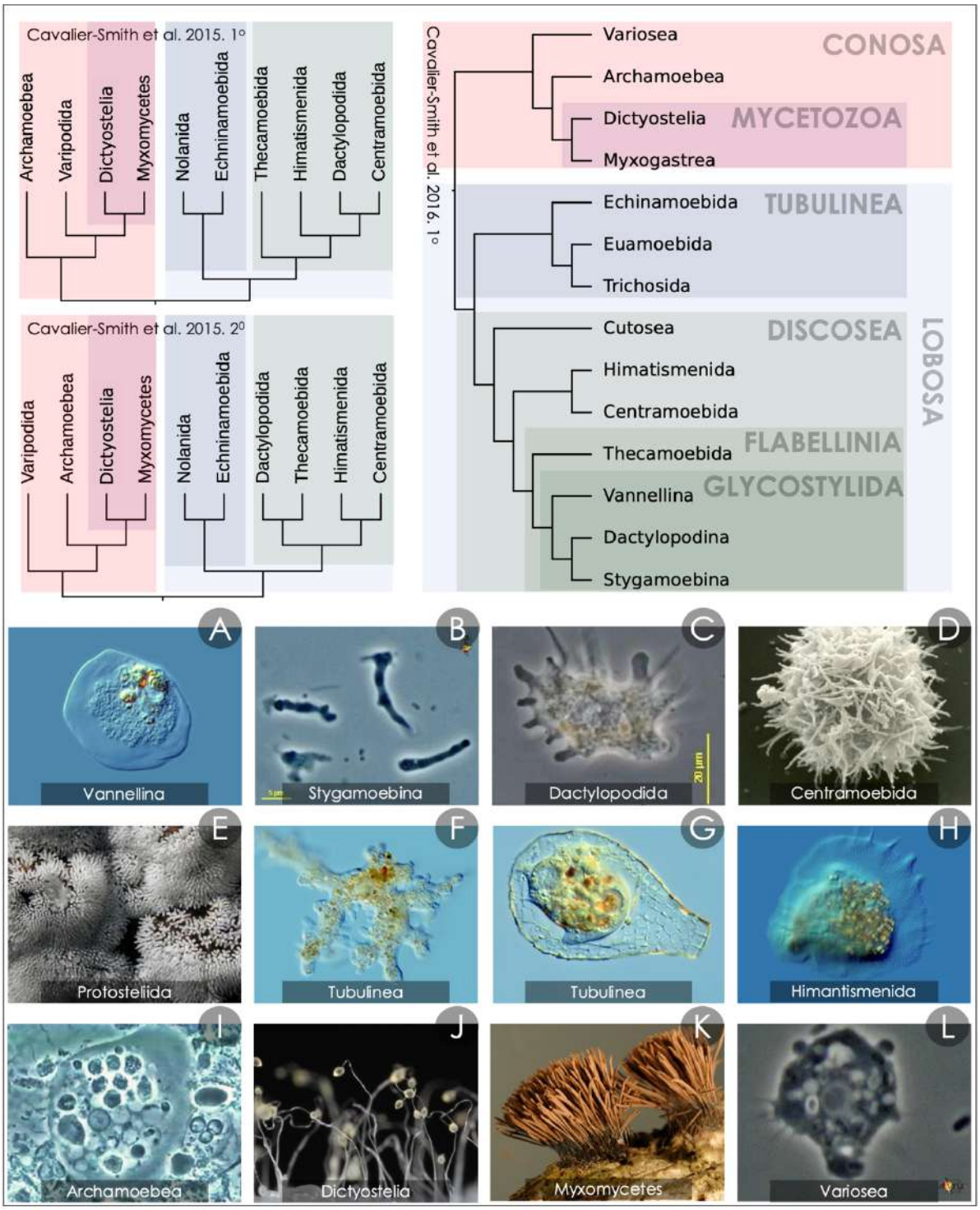


Figure 6. Phylogenetic diversity of amoebozoa after Cavalier-Smith et al. (2015) and Cavalier-Smith et al. (2016) and a photographic glimpse into the morphological diversity of amoebozoa. A. ***Vannella simplex*** (source EoL, author Projecto Agua Flickr Group, CC-BY-NC-SA); **B.** ***Stygamoeba*** **sp**. (source EoL, author micro*scope, licence CC-BY-NC); **C.** ***Paramoeba*** **sp**. (source EoL, author micro*scope, licence CC-BY-NC); **D.** ***Acanthamoeba castellanii*** (source EoL, author Katz Lab Flickr Group, CC-BY NC 2.0); **E.** ***Ceratiomya fruticulosa*** (source EoL, author BioImages, the virtual fieldguide UK, CC-BY-NC-SA 3.0); **F.** ***Amoeba proteus*** (source EoL, author Projecto Agua Flickr Group, CC-BY-NC-SA): **G.** ***Quadrulella symmetrica*** (source EoL, author Projecto Agua Flickr Group, CC-BY-NC-SA); **H.** ***Cochliopodium*** **sp.** (source EoL, author Projecto Agua Flickr Group, CC-BY-NC-SA**); I.** ***Entamoeba gingivalis*** (source EoL, author Institut international parodontie, licence CC-BY 3.0); **J.** ***Dictyostelium discoideum*** (source EoL, author Usman Bashir, CC-BY-SA 3.0); **K.** ***Stemonitis axifera*** (source EoL, author Dr. Lorne Stobbs (Stobbsl), licence CC-BY 3.0**); L.** ***Filamoeba*** **sp**. (source EoL, author micro*scope, licence CC-BY-NC). Relationships within Amoebozoa are not well understood yet, and new high-ranked taxa are being discovered all the time.

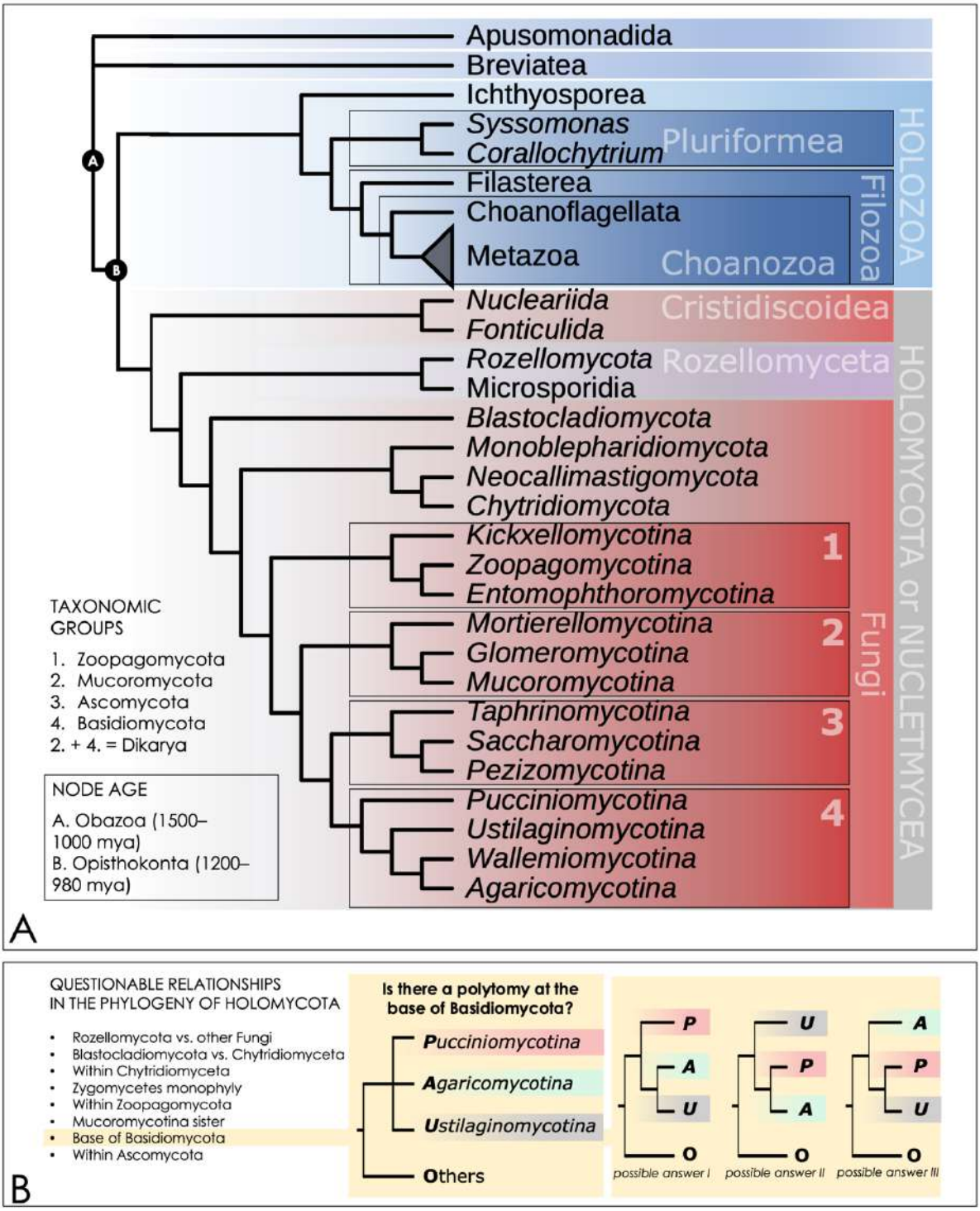


Figure 7. A glimpse into the phylogenetic diversity of Obazoans (Breviates, Apusomonadids, and Opisthokonts). Obazoans include animals, plants and their less-known microscopic relatives, such as microsporidians, ichthyosporeans, and cristidiscoideans. Many open questions still exist and relationships within those groups have not been clearly resolved yet. **A.** Simplified phylogeny of Obazoa, with well-known higher taxonomic groups denoted, as well as certain node ages (in mya – millions of years ago). Reference tree was constructed following results of Adl et al. (2019), Brown et al. (2013), Cavalier-Smith (2013), Dohrmann & Wörheide (2017), Hibbett et al. (2007), Kiss et al. (2019), Li et al. (2021), and Spatafora et al. (2016). **B.** Open questions in fungal evolutionary biology (after Li et al. 2021) and four possible topologies as answers to the question "Is there polytomy at the base of Basidiomycota?", all equally likely (three resolved and an unresolved one, depicted with the question).

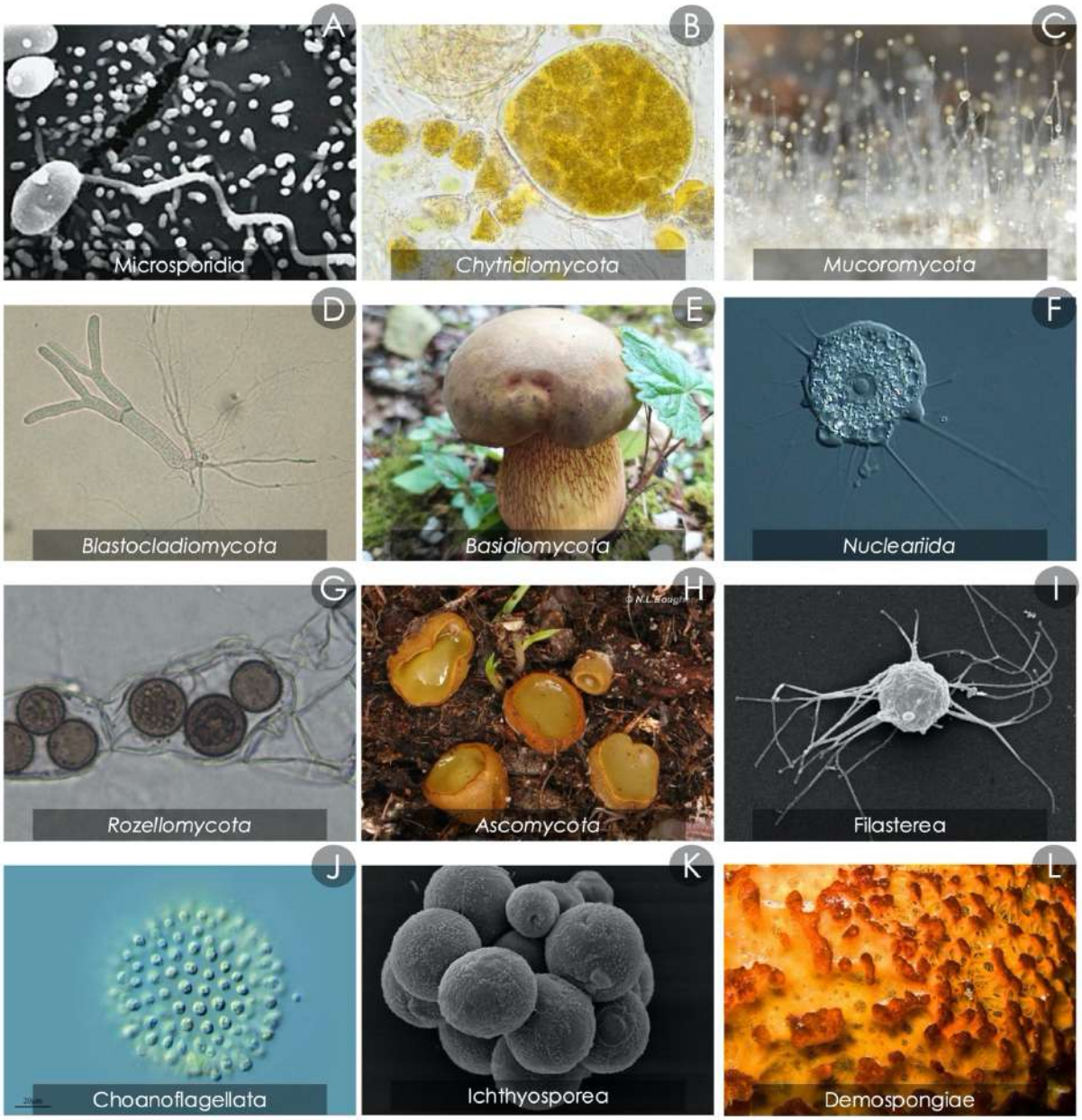


Figure 8. A glimpse into the morphological diversity of Opisthokonta (Nucletmycea or Holomycota, and Holozoa). *Italicized are taxa covered by the International Code of Nomenclature for algae, fungi, and plants (Turland et al. 2018).* **A. microsporidian spore** (source Wikimedia commons, Centers for Disease Control and Prevention, Public domain); **B. *Synchytrium erieum*** (source EoL, author Bioimages, licence CC-BY-NC-SA 3.0); **C. *Mucor mucedo*** (source Wikimedia commons, author James Lindsey, CC-BY-SA 3.0); **D. *Allomyces* sp.** (source Wikimedia commons, (author TelosCricket, licence CC-BY SA 4.0); **E. *Suillelus* sp**., (photo J. Skejo, Trakošćan, Croatia); **F. *Linderina pennispora***, source EoL, author Connierobertson3, licence CC-BY-NC 2.0; **F. *Nucleariа thermophila*** (source Wikimedia commons, author ja:User:NEON / User:NEON_ja, licence CC BY-SA 2.5);**G. *Rozella allomycis*** (source Wikimedia commons, author Timothy James derivative work: Toter Alter Mann, licence CC BY-SA 3.0); **H. *Aleurina ferruginea* (**source Wikimedia commons, author N. L. Boulgher, Public domain); **I. *Ministeria vibrans*** (source EoL, author Iaki Ruiz-Trillo CC-BY-NC 2.0); **J. *Sphaerotheca* sp**. (source EoL, author Proyecto Agua Flickr Group, CC-BY-NC-SA); **K. *Sphaeroforma arctica*** (source EoL, author Iaki Ruiz-Trillo CC-BY-NC 2.0); **L. *Tethya californiana*** (source EoL, author Ken-ichi Ueda, CC-BY-NC 4.0).

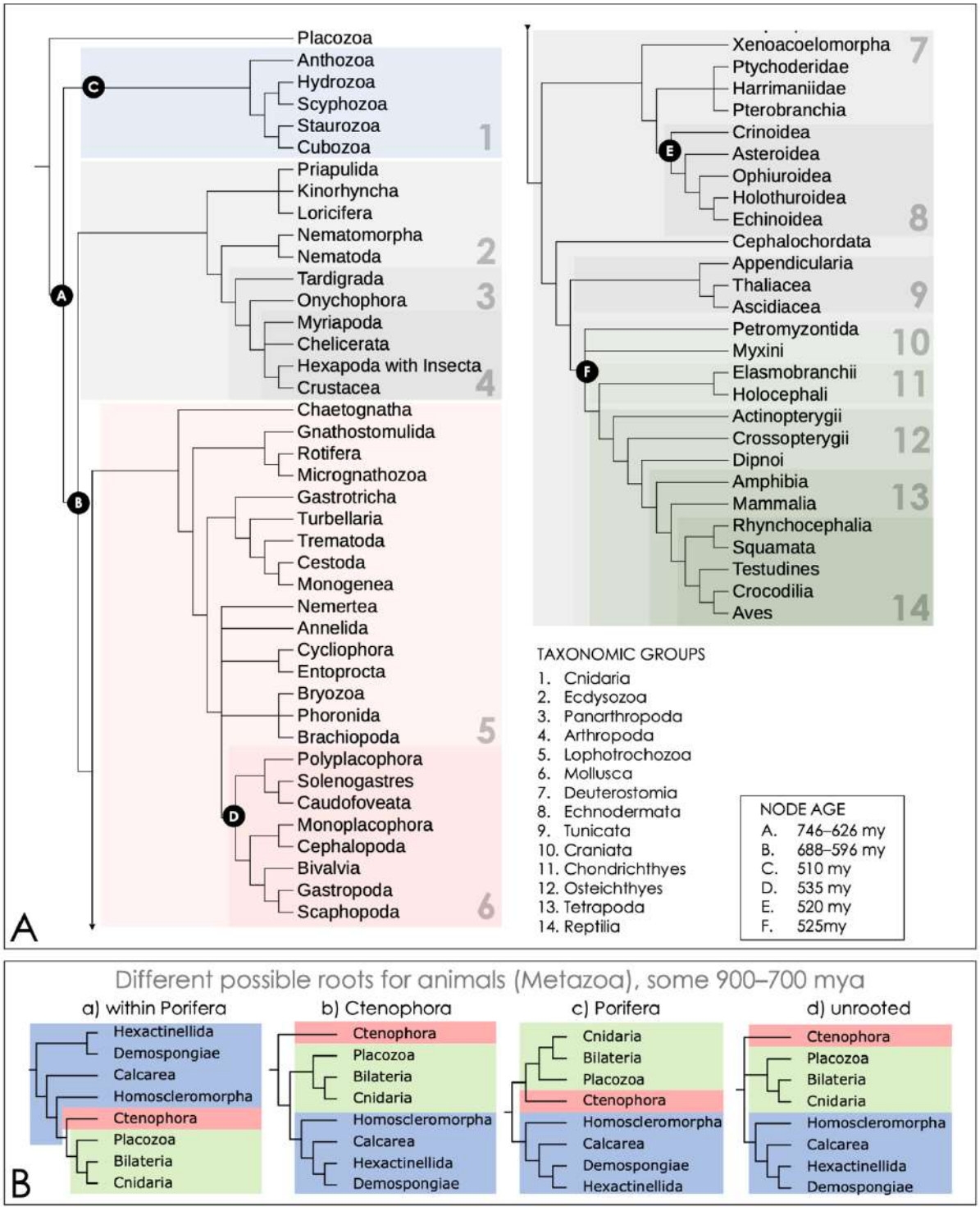


Figure 9. A glimpse into phylogenetic diversity of animals (Metazoa). A. Simplified phylogeny of Eumetazoa (Placozoa, Bilateria, and Cnidaria), with well-known higher taxonomic groups denoted, as well as certain node ages (in Mya – millions of years ago). B. Possible positions of the metazoan root. Reference trees constructed following Zardoya et al. (1998), Nielsen (2012), Nosenko et al. (2013), Glenner et al. (2014), Borowiec et al. (2015), Whelan et al. (2015), Simion et al. (2017), Laumer et al. (2019), Dohrmann & Wörheide (2017).

Figure 10. A glimpse into morphological diversity of eumetazoan animals. Animals are the most specious eukaryotic group, with more than million described species. Most of those species belong to Arthropoda, more specifically Insecta, and more specifically Holometaboa, i.e., the most diverse living beings on the planets are the insects with larva and coccon. **A.** Actiniaria, Mljet, Croatia. **B.** *Propithecus coquereli* A. Grandidier, 1867, Madagascar. **C.** *Mantella aurantiaca* Mocquard, 1900, Madagascar. **D.** *Ocypode ceratophthalmus* (Pallas, 1772), Madagascar. **E.** *Gasteracantha versicolor* (Walckenaer 1841), Madagascar. **F.** *Glomeris pulchra* C.L.Koch, 1847, Sedramić, Croatia. **G.** *Hermodice carunculata* (Pallas, 1766), Mljet, Croatia. **H**. *Echinaster sepositus* (Retzius, 1783) (Asteroidea) and maybe *Arbacia lixula* (Linnaeus, 1758) (Echinoidea), Mljet, Croatia. **I.** *Meledella werneri* Sturany, 1908, Mljet, Croatia. **J.** *Brookesia therezieni* Brygoo et Domerque, 1970, Madagascar. **K.** *Oecanthus dulcisonans* Gorochov, 1993, Mljet, Croatia. **L.** *Nesoenas picturatus* (Temminck, 1813), Madagascar. Photo J. Skejo.

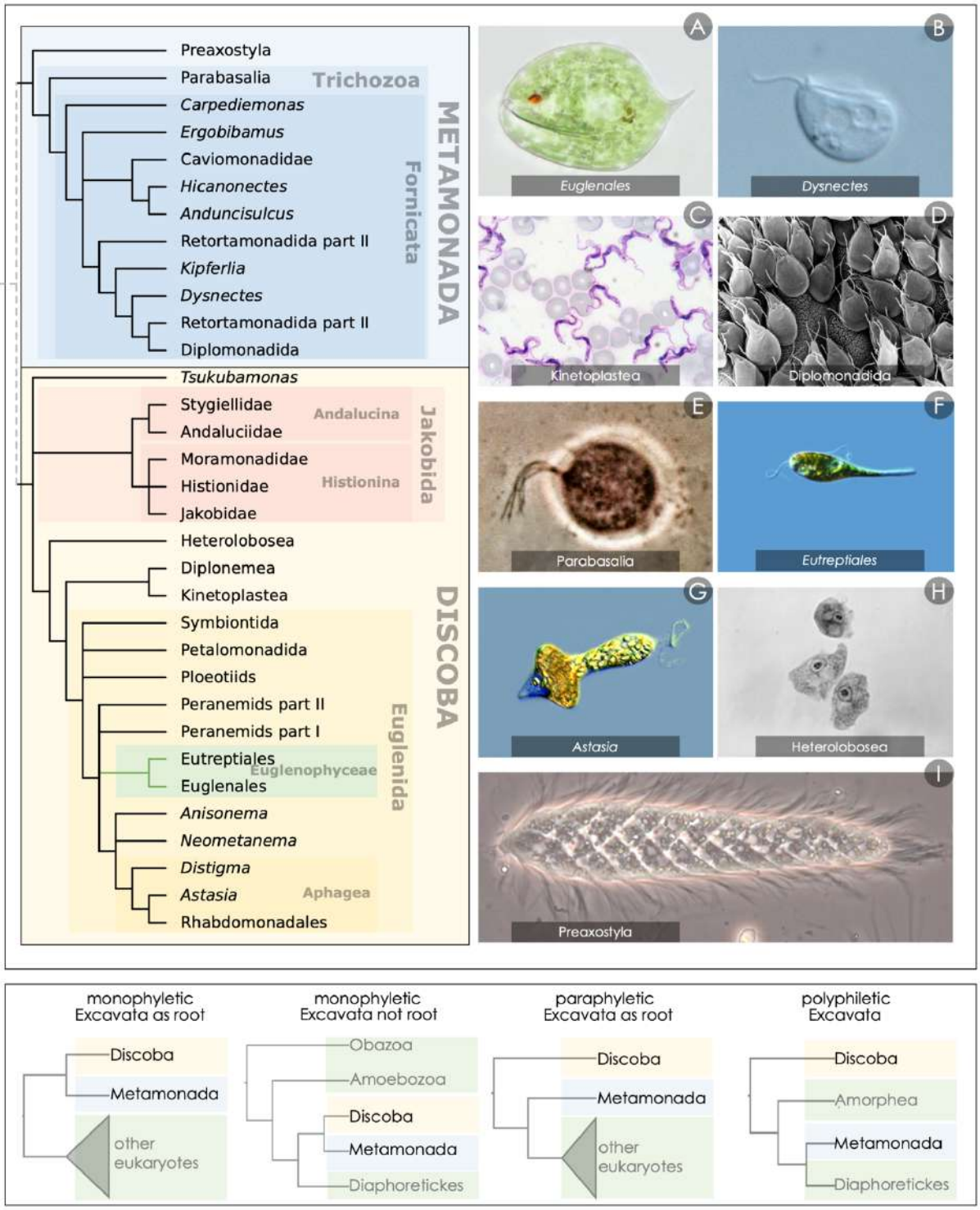


Figure 11. A glimpse into the morphological and the phylogenetic diversity of Excavata. Exavate relationship towards other eukaryotes has not been well-understood yet, hence some of the possible topologies are shown. Phylogenetic relationships between the groups follow Adl et al. (2019), Cavalier-Smith (2016), Leander et al. (2017), Pánek et al. (2015), Kamikawa et al. (2014), Simpson et al. (2017) and Yukubi et al. (2017). Species on the figures: ***A. Phacus triqueter*** (source EoL, author BioImages, the virtual fielguide, UK, CC-BY-NC-SA 3.0); **B. *Dysnectes brevis*** (source EoL, author Naoji Yubuki, CC-BY 3.0); **C. *Trypanosoma* sp.** (source EoL, author Alan R. Walker, CC-BY-SA 3.0); **D. *Giardia duodenalis*** (= G. lamblia) (source EoL, author Public Health Images Library, Public domain); **E. *Trichomonas vaginalis*** (source EoL, author Dr Graham Beards, CC-BY 3.0); **F. *Eutreptia media*** (source EoL, author Proyecto Agua, CC-BY-NC-SA); **G. *Astasia* sp.** (source Eol, author Proyecto Agua, CC-BY-NC-SA); **H. *Naegleria gruberi*** (source Wikimedia commons, author Dr. George Healy, Public domain); **I. *Dinenympha exilis*** (source Wikimedia commons, author Djpmapleferryman, CC BY-SA 4.0) .

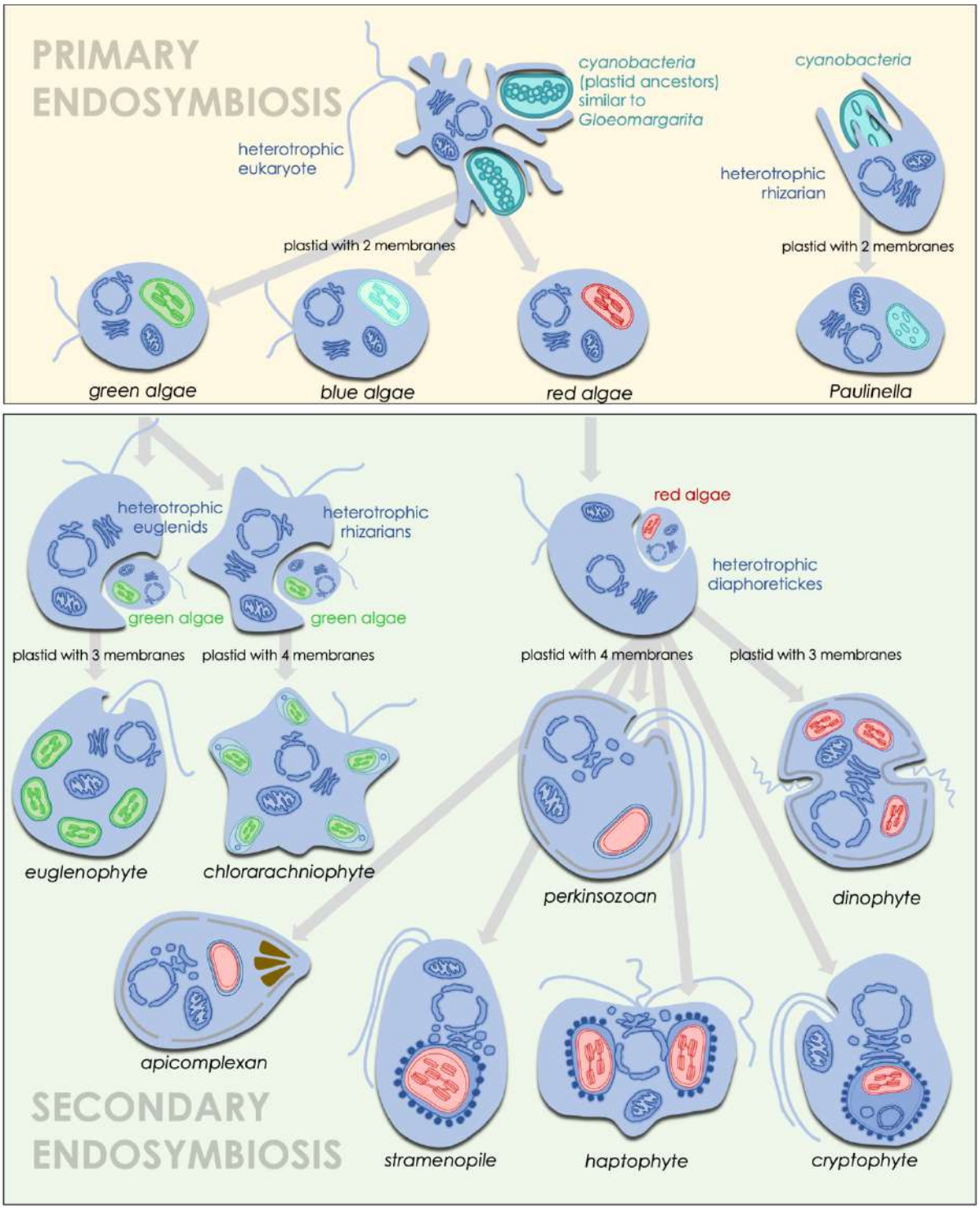


Figure 12. Origin of primary and secondary plastids, a schematic visualization. Plastids found in Archaeplastida (Rhodophyta, Chloroplastida and Glaucophyta) are the primary ones. The primary plastid means that is has two membranes and that it originated via primary endosymbiosis of a heterotrophic eukaryote with a cyanobacterium. Secondary plastids originated via endosymbiosis between a heterotrophic eukaryote and a photosynthetic eukaryote. Because of that, eukaryotes that underwent the secondary endosymbiosis have three of four membranes (e.g., heterokont algae or chrolarachniophyceans). Two of those membranes originate from the original plastid that was inside a photosynthetic eukaryote, while third and/or fourth membrane belonged to the endosymbiotic eukaryote, and eukaryotic host, respectively. There are new examples of primary endosymbiosis that is currently happening, for example in the genus *Paulinella*. The schematic origin follows Gould et al. (2008).

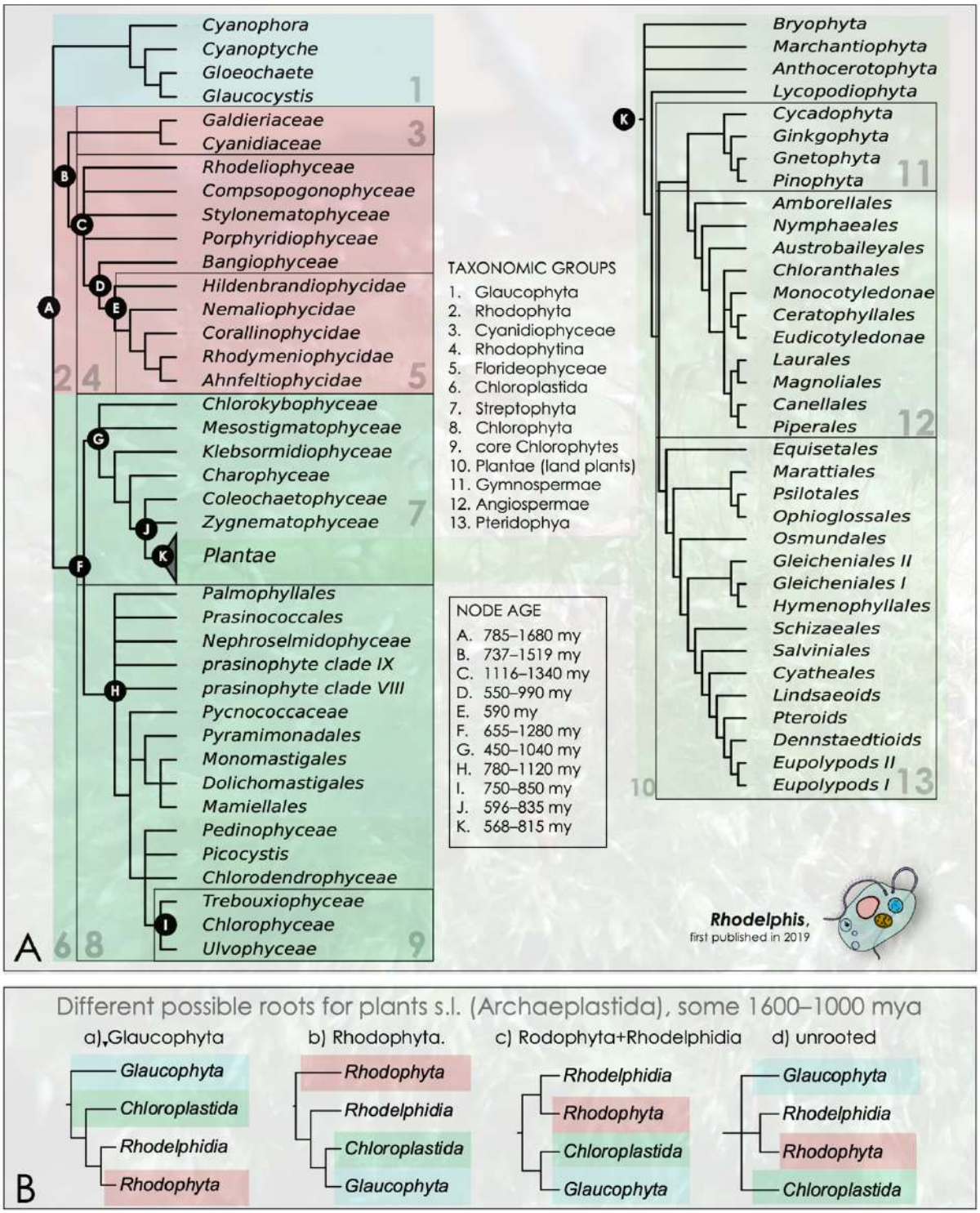


Figure 13. A glimpse into phylogenetic diversity of plants (Archaeplastida). A. Simplified phylogeny of Glaucophyta, Chloroplastida, Rhodelphidia, and Rhodophyta), with well-known higher taxonomic groups denoted, as well as certain node ages (in mya – millions of years ago). B. Possible positions of the archaeplastidian root. Reference trees constructed following De Clerck et al. (2012), Shen et al. (2018), Willis & McElwain (2014), APG IV (2016), Colp & Archibald (2019), and de Vries & Gould (2018).

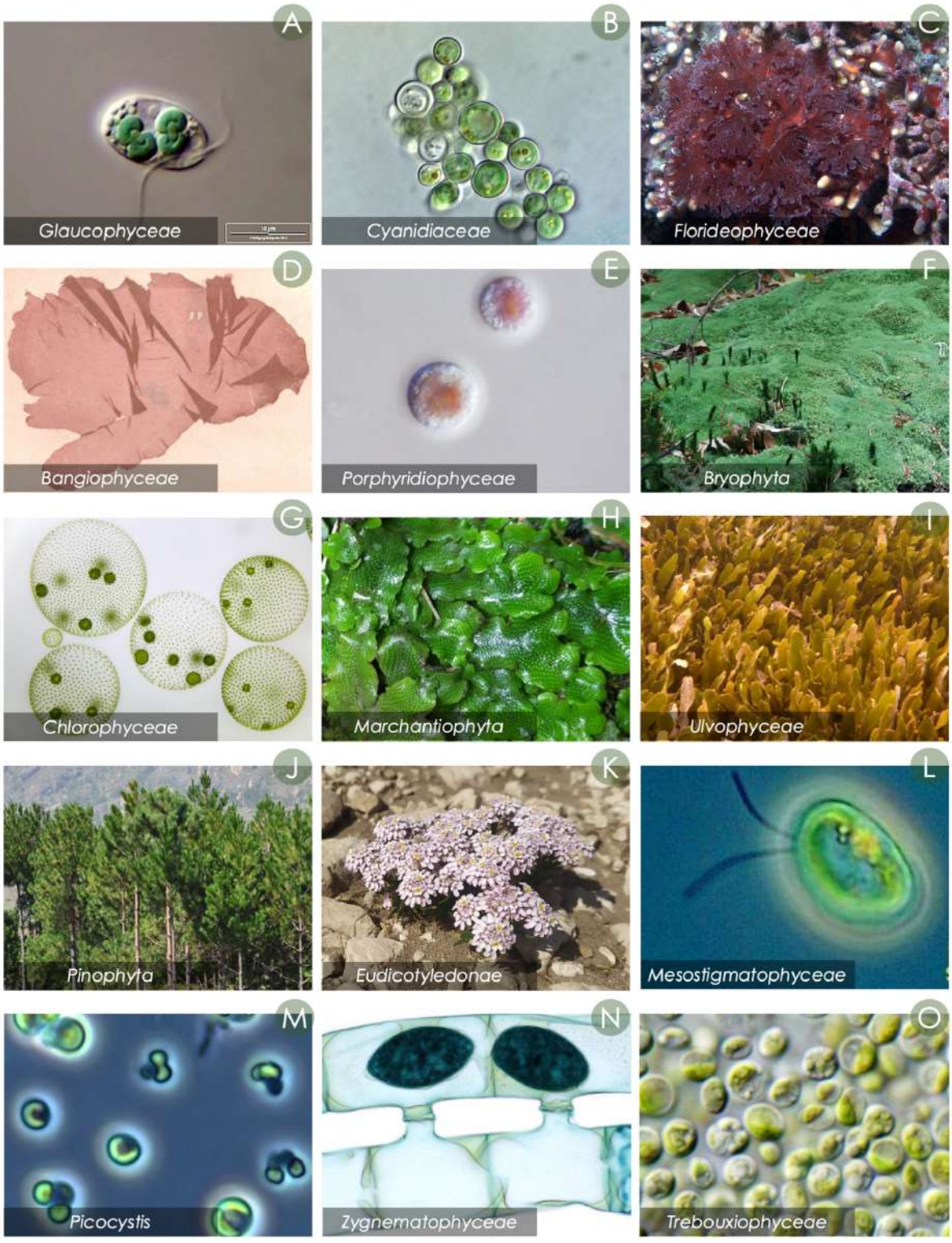


Figure 14. A glimpse into the morphological diversity of Archaeplastida, i.e., plants (*sensu latissimo*). A. *Cyanophora paradoxa* (Wolfgang Bettighofer at www.protisten.de, Plankton Net (Biodiversity Data Provider), CC BY-SA 3.0); **B. *Cyanidium* sp**. (Wikimedia, User:NEON_ja, CC BY-SA 3.0); **C. *Peyssonnelia squamaria (***Guido Picchetti - http://www.guidopicchetti.it, CC BY-SA 3.0); **D. *Porphyra* sp**. (Emile Wuitner, Public domain); **E. *Porphyridium purpureum (***Wikimedia, Nebodo, CC BY-SA 4.0); **F. *Leucobryum glaucum***, J. Skejo, Trakošćan, Croatia; **G. *Volvox* sp.** (Wikimedia, Frank Fox, CC BY-SA 3.0 de.); **H. *Conocephalum conicum***, J. Skejo, Ivanščica Mt., Croatia; I. ***Caulerpa profilera*** (Wikimedia, Nanosanchez, Public domain); **J. *Pinus kesiya***, J. Skejo, Madagascar; **K. *Iberis saxatilis***, J. Skejo, Dinara Mt., Croatia; **L. *Mesostigma viride (***Encyclopaedia of Life, David Patterson and Bob Andersen, CC-BY NC); **M. *Picocystis* sp.** (Microbial Life, Carleton University. serc.carleton.edu/microbelife/topics/monolake/general.html, CC0); **N. *Spirogyra* sp.** (Wikimedia, Jon Houseman & Matthew Ford, CC BY-SA 4.0); **O. *Chlorella vulgaris***, (Wikimedia, / User:NEON_ja, CC BY-SA 3.0).

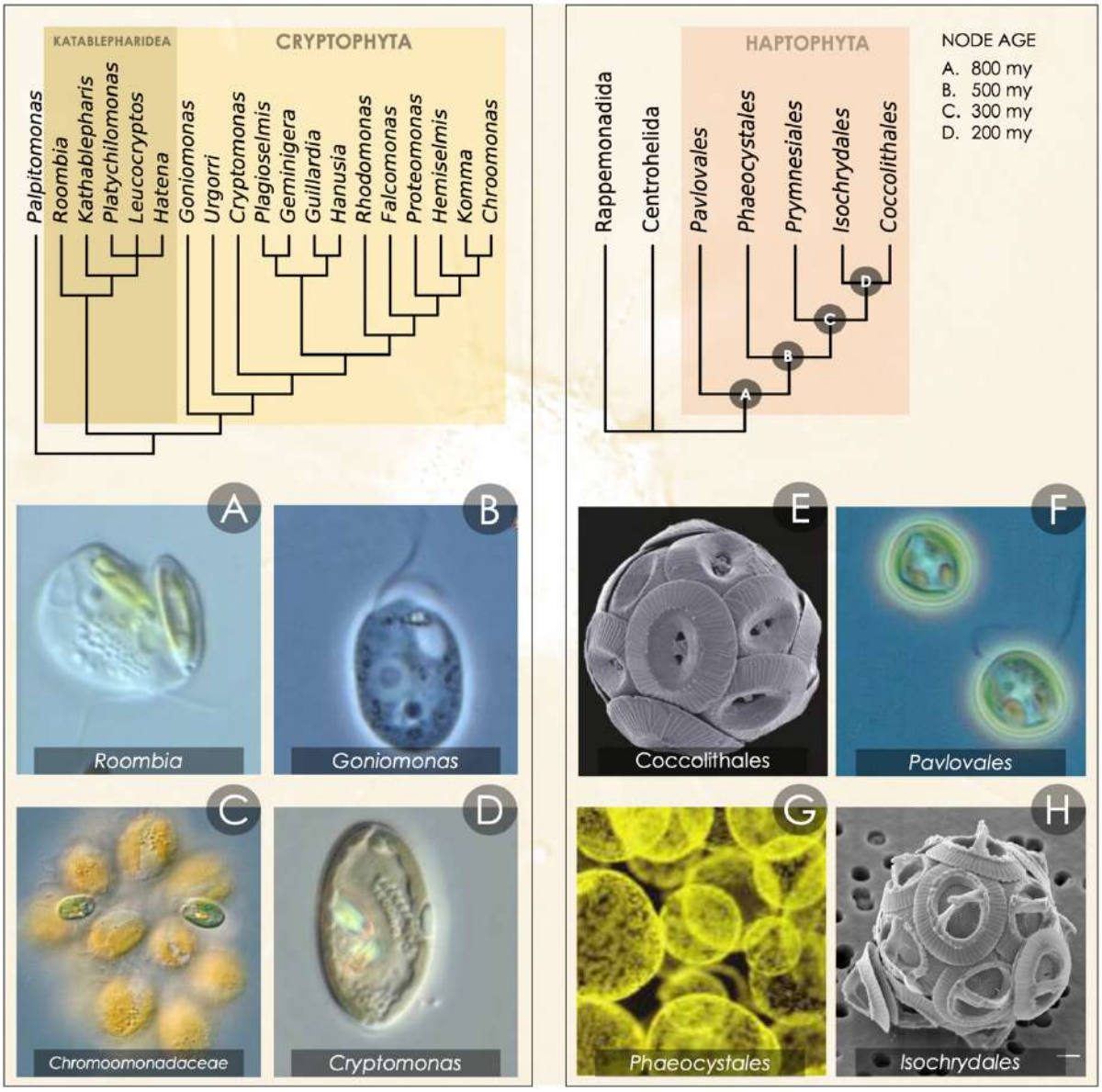


Figure 15. An insight into the phylogenetic and morphological diversity of Hacrobia – Haptista, and Cryptista. Phylogenetic relationships follow Adl et al. (2019), Eikrem et al. (2017), and Hoef-Emden & Archibald (2017). A. *Roombia truncata* (source EoL, authors Okamoto N, Chantangsi C, Horak A, Leander B, Keeling P, licence CC-BY 3.0); **B. *Goniomonas* sp.** (source EoL, author micro*scope, licence CC-BY NC); **C. *Chromoomonadaceae*** (source EoL, author Wolfgang Bettighofer, micro*scope, licence CC.BY-NC); **D. *Cryptomonas* sp.** (source EoL, author micro*scope, licence CC-BY NC); **E. *Coccolithus pelagicus* ssp. *braarudii***, fossil species (source Wikmedia Commons, authors Richard Lampitt, Jeremy Young, The Natural History Museum, London, CC BY 2.5 licence); **F. *Pavlova* sp.** (source EoL, micro*scope, licence CC.BY-NC); **G. *Phaeocystis antarctica*** colonies, source Wikimedia commons, authors Bender et al. (2018), CC BY 4.0; **H. *Gephyrocapsa oceanica*** (source EoL, author ja:User:Neon/commons:User:Neon_ja, CC-BY-SA 3.0). In the background there is a re-colored photo of ***Raphidiophrys contractilis*** (Centrohelida) (by ja:User:NEON / commons:User:NEON_ja - Own work, CC BY-SA 2.5).

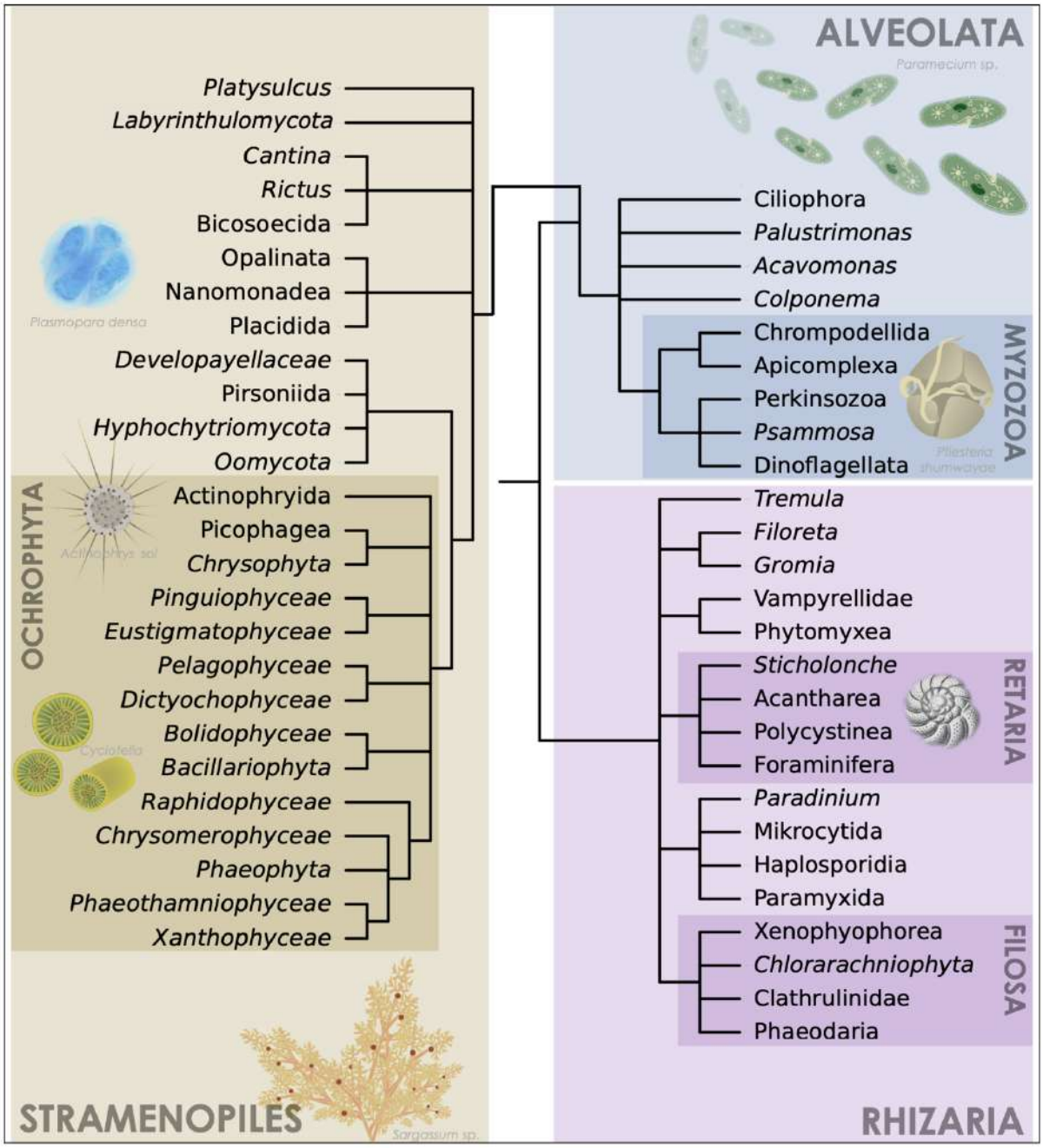


Figure 16. A glimpse into the phylogenetic diversity of SAR – Stramenopiles, Alveolata, and Rhizaria. Many small groups whose relationship to others is not understood are not shown. Relationships follow Simpson et al. (2017), drawn after Bass et al. (2009), Riisberg et al. (2009), Cavalier-Smith & Scoble (2013), Tikhonenkov et al. (2014), Janouškovec et al. (2015), Park & Simpson (2015), Shiratori et al. (2015), Yubuki et al. (2015), Burki et al. (2016), Derelle et al. (2016), Sierra et al. (2013, 2016), and Krabberod et al. (2017). Small images show SAR examples: ***Paramecium* sp**. (Ciliophora), ***Cibicides pachyderma*** (Foraminifera), ***Plasmopara densa*** (*Oomycetes*) (source EoL, author BioImages, the virtual fieldguide, UK, CC-BY-NC-SA 3.0); ***Pfiesteria shumwayae*** (Dinophlagellata), ***Sargassum* sp**. (Phaeophyta), ***Cyclotella* sp**. (Bacillariophyta) (Tracey Saxby, Integration and Application Network, CC BY-SA 4.0), ***Actinophrys sol*** (Actinophryida) (Kim Kraeer, Lucy Van Essen-Fishman, CC BY-SA 4.0).

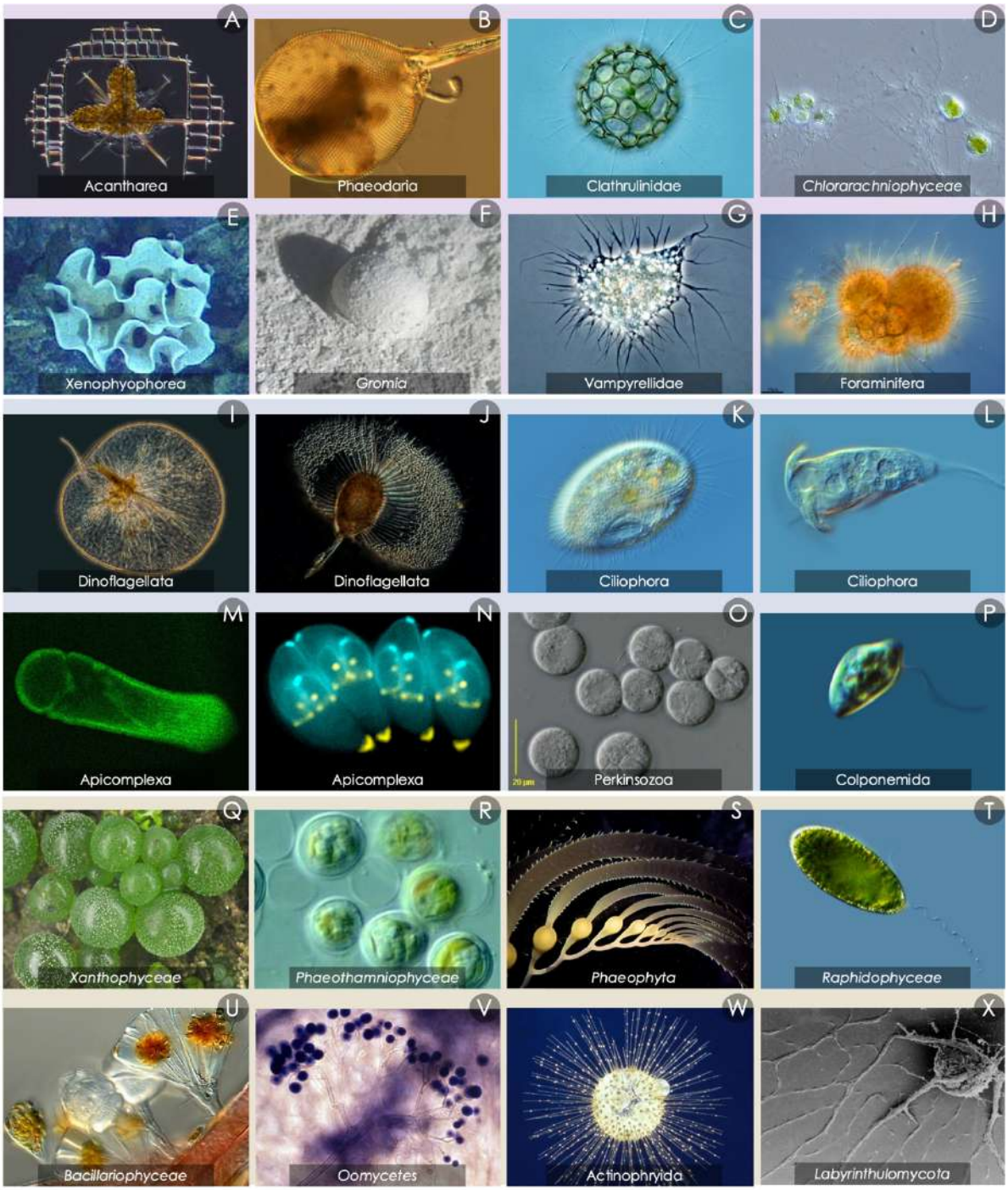


Figure 17. An insight into the morphological diversity of SAR – Rhizaria (A-H); Alveolata (I-P), and Stramenopiles (Q-Z). A. *Lithoptera fenestrata* (EoL, John Dolan, CC-BY-SA 3.0); **B. *Protocystis xiphodon*** (EoL, John Dolan, CC-BY-SA 3.0); **C. *Clathrulina elegans*** (Proyecto Agua, CC-BY-NC-SA); **D. *Chlorarachnion reptans*** (Wikimedia, commons:User:NEON_ja, CC BY-SA 2.5); **E. Xenophyophorea** (NOAA, Public Domain); **F. *Gromia sphaerica*** (Mikhail Matz, Public Domain); **G. *Lateromyxa gallica*** (EoL, Nortbert Hlsmann, CC-BY-NC-SA); **H. *Globigerina bulloides*** (Preyecto Agua, CC-BY-NC-SA); **I. *Noctiluca scintillans*** (Proyecto Agua, CC-BY-NC-SA; **J. *Ornithocerus splendidus*** (EoL, tintinnidguy Flickr Group, CC-BY); **K. *Histobalantium natans*** (Proyecto Agua Flickr Group, CC-BY-NC-SA); **L. *Vorticella* sp.** (Proyecto Agua Flickr Group, CC-BY-NC-SA); **M. *Gregarina garnhami*** (EoL, Garnhami, CC-BY-SA 3.0); **N. *Toxoplasma gondii*** (Wikimedia, authors Ke Hu and John M. Murray, CC-BY-SA 4.0); **O. *Perkinsus* sp.** (EoL, David Patterson, Linda Amaral Zettler, Mike Peglar and Tom Nerad, micro*scope, CC-BY-NC); **P. *Oxyrrhis* sp.** (Proyecto Agua Flickr Group, CC-BY-NC-SA); **Q. *Botrydium granulatum*** (EoL, BioImages, the virtual field guide, UK, licence CC-BY-NC-SA 3.0); **R. *Phaeoschizochlamys* sp.** (EoL, David Patterson and Bob Andersen, micro*scope, CC-BY-NC); **S. *Macrocystis pyrifera*** (EoL, California Academy of Sciences, CC-BY-NC-SA 3.0); **T. *Vacuolaria virescens*** (EoL, Proyecto Agua Flickr Group, CC-BY-NC-SA); **U. *Licmophora juergensii*** (EoL, Wolfgang Bettighofer, CC-BY-NC-SA); **V. *Hyaloperonospora parasitica*** (EoL, Emmanuel Boutet, CC-BY-SA 3.0); **W. *Actinophrys sol*** (EoL, Amgueddfa Cymru - National Museum Wales Flickr Group, CC-BY-NC-SA); **X. *Labyrinthula coenocystis*** (EoL, Norbert Hlsmann, CC-BY-NC-SA).

HYPOTHESES AND AIMS

Eleven important **hypotheses**, some of which are mutually compatible, that have not been rejected so far and that should be taken into account when studying eukaryotic origin and the eukaryotic common ancestor are as follows. **i)** Eukaryogenesis marks transition between FECA (the First Eukaryotic Common Ancestor) and LECA (the Last Eukaryotic Common Ancestor) (O'Malley et al. 2019). **ii)** Eukaryotes are monophyletic group, and their ancestor is LECA (the Last Eukaryotic Common Ancestor) (Martin & Müller 1998). **iii)** LECA originated via symbiogenesis between archaea (close to *Prometheoarchaeum syntrophicum*) and proteobacteria (alphaproteobacteria or closely related group) (Imachi et al. 2020; Fan et al. 2020; Muñoz-Gómez et al. 2022). **iv)** The eukaryotic genome is chimaera composed of archaeal and bacterial genes (Brückner & Martin 2020). **v)** Proteobacteria survived during eukaryogenesis as mitochondria inside the host's cytosol (Martin & Müller 1998, Roger et al. 2017). **vi)** The mitochondrial genome is purely bacterial (Gray et al. 2012; Roger et al. 2017). **vii)** LECA had a complex endomembrane system, composed of nucleus, endoplasmic reticulum, and Golgi apparatus (Garg et al. 2016; Gould et al. 2016). **viii)** LECA was heterotop**h**ic (Martin et al. 2001). **ix)** Primary plastids and autotrophic eukaryotes originated via cyanobacterial endosymbiosis (Ponce-Toledo et al. 2017). **x)** LECA did not have plastids; they instead came came a bit later. **xi)** LECA was sexual, but but some of its descendants lost sex (Hofstatter & Lahr 2019).

The general aim of this thesis is reconstruction of the Last Eukaryotic Common Ancestor (LECA) from several different angles and by the means of cladistic and phylogenetic analyses. The **specific aims** of the thesis are:

(1) to test whether paraphyletic groups represent monophyletic groups that can be used in modern evolutionary biology, and more specifically, in eukaryotic classification. (Chapter **Paper 1: Skejo & Franjević 2020**).

(2) to test by annotating genome duplications if the mitochondrion was involved early in the process of eukaryogenesis, i.e., to see how many mitochondria-derived genes LECA had. (Chapter **Paper 2: Tria et al. 2021**).

(3) to test by the means of ancestral state reconstruction if LECA was a multinucleate (syncytial/coenocytic) organism who exhibited closed nuclear division.
(Chapter Paper 3: **Skejo et al. 2021**).

PAPERS

I.

SKEJO & FRANJEVIĆ (2020)

EUKARYOTES ARE A HOLOPHYLETIC GROUP OF POLYPHYLETIC ORIGIN.

FRONTIERS IN MICROBIOLOGY.

II.

TRIA ET AL. (2021)

GENE DUPLICATIONS TRACE MITOCHONDRIA TO THE ONSET OF EUKARYOTE COMPLEXITY.

GENOME BIOLOGY AND EVOLUTION.

III.

SKEJO ET AL. (2021)

EVIDENCE FOR A SYNCYTIAL ORIGIN OF EUKARYOTES FROM ANCESTRAL STATE RECONSTRUCTION.

GENOME BIOLOGY AND EVOLUTION.

frontiers in Microbiology

OPINION
published: 02 July 2020
doi: 10.3389/fmicb.2020.01380

Eukaryotes Are a Holophyletic Group of Polyphyletic Origin

Josip Skejo [1,2] *and Damjan Franjević* [2]*

[1] *Institute of Molecular Evolution, Heinrich–Heine University Düsseldorf, Düsseldorf, Germany,* [2] *Evolution Lab, Division of Zoology, Department of Biology, Faculty of Science, University of Zagreb, Zagreb, Croatia*

Keywords: Eukaryomorpha, archaea, alphaproteobacteria, eukaryogenesis, lichens, hybridization, symbiogenesis, paraphyly

INTRODUCTION

All living beings can be assigned to one of the three domains of life (Woese et al., 1990; Williams et al., 2013), all of which are monophyletic (Doolittle, 2014). Two prokaryotic domains, Archaea and Bacteria, are characterized by the lack of intercellular compartments (Martin, 1999; McInerney et al., 2014), whereas eukaryotes, characterized by the complexity of cellular structures and life cycle, originated via symbiogenesis of an archaeal host and a bacterial endosymbiont i.e. proto-mitochondrion (Mereschkowsky, 1905; Zimorski et al., 2014; Muñoz-Gómez et al., 2017; Roger et al., 2017). With millions of described species (Costello et al., 2013; Adl et al., 2019), eukaryotes are morphologically the most diverse of the three groups bearing symbiogenesis as the hallmark of their evolutionary origin (Wallin, 1927; Margulis, 1991). Symbiogenesis has always been a common phenomenon in the eukaryotic evolution (McFadden, 2001; Nowack and Melkonian, 2010; Bonfante and Desirò, 2017). Nevertheless, there are still many unanswered questions regarding the prokaryotes that participated in eukaryogenesis. The true evolutionary position of eukaryotes is hence the subject of continuing debates and it has still not been widely agreed if eukaryotes represent a separate domain (Williams et al., 2013; Doolittle, 2020). Alphaproteobacteria is known to be the ancestor of mitochondria (Roger et al., 2017). However, our understanding of the archaeal lineage that gave rise to the eukaryotic nuclear genome is still insufficient. Asgard archaea, which were recently identified based on metagenome-assembled sequences (Spang et al., 2015; Seitz et al., 2016; Zaremba-Niedzwiedzka et al., 2017; MacLeod et al., 2019), possess eukaryotic signature proteins (ESPs) involved in cytoskeleton regulation (Akil and Robinson, 2018; Akil et al., 2019), and are being cultivated now (Imachi et al., 2020). The first photographed member of Asgard is known under the name "Candidatus *Prometheoarchaeum syntrophicum,*" and it does not exhibit eukaryotic features (such as the presence of mitochondrion, nucleus, endoplasmic reticulum, or sexual reproduction), but rather exhibits typical prokaryotic features, such as small size, spherical (coccoid) body, and lack of organelles (Imachi et al., 2020). Recently, Fournier and Poole (2018) presented a taxonomic view in which Asgard represented the main eukaryotic ancestor (parent) and were, along with eukaryotes, united into a "monophyletic" group named Eukaryomorpha. The aim of this opinion manuscript is to debate this newly introduced term. We briefly review the meaning of the terms "monophyletic" and "polyphyletic," and we draw attention to the bacterial contribution to eukaryogenesis.

PARAPHYLETIC MEANS MONOPHYLETIC

Evolutionary biologists use the term "monophyly" in various ways (see e.g., Envall, 2008), just as Hennig (1950, 1966), the creator of the term, originally did, which has hitherto ensued a lot of confusion (Envall, 2008). In this opinion, we use the term "monophyletic" only for groups with a single definable ancestor, meaning that paraphyletic groups are also considered

OPEN ACCESS

Edited by:
Andrés Moya,
University of Valencia, Spain

Reviewed by:
Luis Delaye,
Instituto Politécnico Nacional de México (CINVESTAV), Mexico
Daniel Tamarit,
Wageningen University and Research, Netherlands

***Correspondence:**
Damjan Franjević
damianf@biol.pmf.hr

Specialty section:
This article was submitted to Evolutionary and Genomic Microbiology, a section of the journal Frontiers in Microbiology

Received: *03 March 2020*
Accepted: *28 May 2020*
Published: *02 July 2020*

Citation:
Skejo J and Franjević D (2020) Eukaryotes Are a Holophyletic Group of Polyphyletic Origin. Front. Microbiol. 11:1380. doi: 10.3389/fmicb.2020.01380

as monophyletic. Each taxonomic group can be characterized by either having a shared (single) ancestor—"monophyletic group" *or* having numerous ancestors—"polyphyletic group" (Hennig, 1950, 1966). Polyphyletic groups are not taxonomically desirable, and traditionally, characters shared by members of such a group represent *homoplasies* (analogies), i.e., traits that evolved independently in similar environments on account of similar selective advantages (Wake et al., 2011). A historical error occurred when Hennig (1950, 1966) defined two groups, monophyletic and paraphyletic, based on the inclusion of all descendants of a given ancestor. If all the descendants of a given ancestor belonged to one group, it was regarded as a *monophyletic* group, and if this was not the case, it was regarded as a *paraphyletic* group (Hennig, 1950, 1966). Missing from such definition was the distinction between a group with a single ancestor and a group that includes all the descendants of an ancestor, which were both defined as monophyletic by Hennig (1950, 1966). Ashlock (1971, 1972, 1974, 1979) noticed the erratum and introduced the term "holophyletic group," referring to a monophyletic group that includes all the descendants of an ancestor. Therefore, a "paraphyletic group" *is* a monophyletic one that does not include all the descendants of an ancestor (**Figures 1A–C**).

Well-known examples of holophyletic groups are mammals (descendants of Therapsida), snakes (descendants of earless and legless lizards), birds (descendants of Dinosauria), modern amphibians, tetrapods (land vertebrates, descendants of fish), jawed vertebrates, bilaterians (bilaterally symmetric animals), animals, and eukaryotes (Pough et al., 1999; Nielsen, 2012; Doolittle, 2014). Examples of paraphyletic groups are reptiles or amniotes (whose descendants are mammals and birds), amphibians (a group including Lissamphibia and extinct amphibians whose descendants are reptiles), sarcopterygians (whose descendants are tetrapods), fish (Pisces) (as they include all vertebrates excluding those inhabiting land), jawless fish (lampreys, hagfish, and extinct groups related to them, whose descendants are also jawed fish), bryophytes in wider sense (as land plants are their descendants), streptophytes (stonewort and relatives, if plants are excluded), archaeplastids (as secondary plastids of SAR and euglenoids are not considered to be archaeplastid members anymore), cyanobacteria (because plastids are regarded as organelles, not cyanobacteria anymore), prokaryotes (because eukaryotes are excluded), Archaea (because the nucleus is not regarded to be an archeon anymore), and Bacteria (because mitochondria are not regarded as Alphaproteobacteria anymore; Pough et al., 1999; Nielsen, 2012; Doolittle, 2014).

If we ignore the presence of mitochondria and existence of lateral gene transfer from bacteria to the eukaryotic host, the origin of the eukaryotic nucleus could be compared to the origin of mammals and birds within amniotes, as described in Fournier and Poole (2018). However, the origin of eukaryotes is not comparable to the origin of these groups, and the bacterial contribution to eukaryogenesis should not be neglected. Eukaryotes are of polyphyletic origin, as their ancestor, LECA, sits on both branches of life—the archaeal (Asgard) and the bacterial branch (Alphaproteobacteria).

POLYPHYLETIC, RETICULATED EVENTS IN EVOLUTION

Well-established examples of natural polyphyletic events include lateral gene transfer (LGT) in prokaryotes (Nelson-Sathi et al., 2015), symbiogenesis in prokaryotes and eukaryotes (biofilms, endosymbiosis, ectosymbiosis, etc.; e.g., Vogels et al., 1980; López et al., 2010; Naumann et al., 2010), and sexual reproduction in eukaryotes (Speijer et al., 2015). Genes can also be of polyphyletic origin; those genes are known as chimeric genes (e.g., Méheust et al., 2018). Polyphyletic origin is an evolutionary event in which two lineages (individuals, populations, or species) merge into a single, "chimeric" lineage. A lineage of polyphyletic origin should not be united with any of its ancestors in an attempt to form a higher monophyletic group, as it will not result in such. Even though eukaryotes are a monophyletic and holophyletic group by definition, they are of polyphyletic origin because of the very nature of their ancestor's, LECA's origin. Today, the polyphyletic origin of eukaryotes is a well-supported scientific theory. Eukaryotic (syn)apomorphies are the traits of eukaryotic complexity: nuclei, mitochondria, Golgi apparatus, endoplasmic reticulum, and sexual reproduction (Koonin, 2010; Koumandou et al., 2013; Garg and Martin, 2016; Doolittle, 2020).

Eukaryogenesis is not a unique example of polyphyletic origin of a monophyletic group. Other such events are widely dispersed in the tree of life. Known examples are hybrid species, which originated via hybridization of two species, usually (but not always) from the same genus (Seehausen, 2004; Grant and Grant, 2008; Meier et al., 2017). *Homo sapiens* is an example of such species. It is a hybrid between *H. heidelbergensis, H. neanderthalensis*, and Denisovians (Sankararaman et al., 2016). The Jutland bow-winged grasshopper (*Chorthippus jutlandica*) is a unique species which originated from the hybridization of *C. brunneus* and *C. biguttulus* in Denmark (Gottsberger, 2007). Domestic wheat is a hybrid between species belonging to the genera *Triticum* and *Aegilops* (Ozkan et al., 2001). There are even examples of one of the ancestral species being extinct, but its mitochondrial genome still being present, which is called a ghost lineage (Recuero et al., 2014). There is no example of a natural monophyletic group that could be composed of any of the aforementioned species and one of its parents, as is the case with Eukaryotes, Asgard, and Eukaryomorpha.

Lichens not only gave rise to the concept of symbiosis (de Bary, 1879), but they are also the classical example of organisms that originated by symbiogenesis (Lutzoni and Miadlikowska, 2009). Lichen species are composed of mycobionts (Ascomycota and/or Basidiomycota) and photobionts (Chlorophyta or Cyanobacteria; Lutzoni and Miadlikowska, 2009; Spribille et al., 2016; Tuovinen et al., 2019). Symbiosis is species-specific (Lindsay, 1856), co-dependent, and the symbionts usually cannot survive outside the lichen. Lichens are an example of a polyphyletic group with multiple polyphyletic origins. Relatives of lichen-forming green algae (symbiont lineages) should not be designated as "Lichenomorpha," even though they represent one of the constituent evolutionary lineages that gave rise to lichens. Cyanobacteria should not be designated as "Plastidomorpha," despite the fact that this group contains the ancestors of plastids.

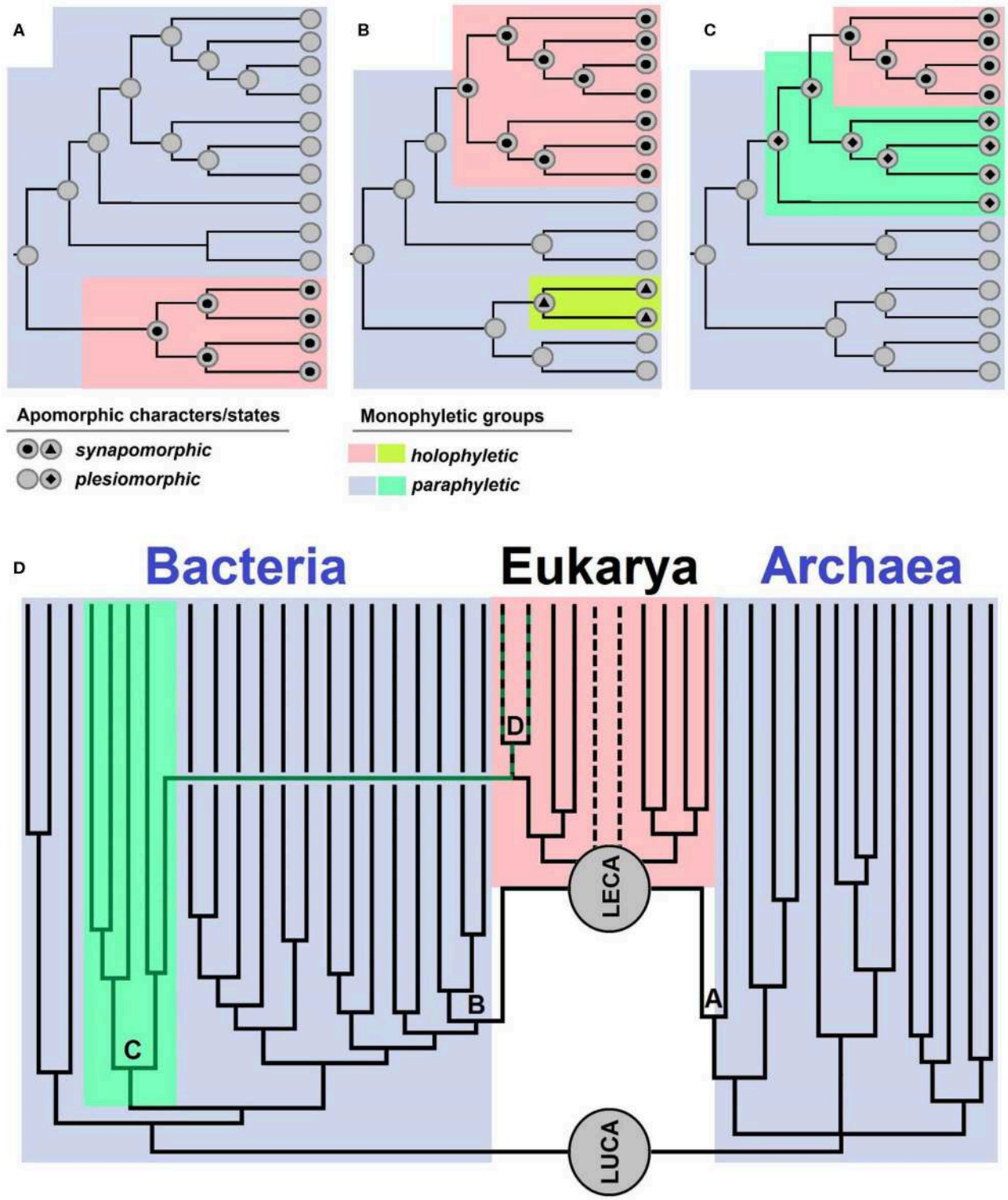


FIGURE 1 | (A–C) Schematic overview of apomorphic characters and states. Synapomorphy is shared by all the members of a group descending from a single ancestor. Plesiomorphy is ancestral and as such not present in all the members of a group. Monophyletic groups are characterized by apomorphies: synapomorphies in holophyletic or plesiomorphies in paraphyletic groups. Topology of the cladograms shown in **(A–C)** is the same, but the distributions of characters and their states are different. Case **(A)** shows a paraphyletic group from which a holophyletic descendant is excluded. Case **(B)** shows a paraphyletic group with two holophyletic

(Continued)

FIGURE 1 | descendants excluded. Case **(C)** shows two paraphyletic groups and a holophyletic group. **(D)** Schematic representation of the evolution of life from its last common ancestor (LUCA), which gave rise to Bacteria and Archaea [the diversity is simplified, and descendants of archeal trichotomy represent Euryarchaeota, TACK+Asgard (Asgard is sister to LECA)]. LECA is the last eukaryotic common ancestor, which originated via a polyphyletic event: symbiogenesis of an archaeon **(A)** which gave rise to nuclei, and Bacteria **(B)**, specifically Alphaproteobacteria, which gave rise to mitochondria. Cyanobacteria **(C)** are a group of bacteria from which the primary plastid **(D)** originated. The dotted lines represent groups with uncertain positions within Eukaryotes.

The case of Archaeplastida (primary photosynthetic eukaryotes) is an interesting one and should be addressed in a separate essay. The supergroup originated via plastidogenesis, an anastomosis between cyanobacteria and eukaryotes; and has since contributed to many anastomoses (secondary endosymbioses) in the eukaryotic tree (McFadden, 2001). The origin of the plastid may be comparable to the origin of mitochondria, however probably only to a certain extent, because of the complexity of the archaeplastidian eukaryotic parent.

BACTERIAL CONTRIBUTION TO EUKARYOGENESIS SHOULD NOT BE NEGLECTED

Bacteria (mainly Alphaproteobacteria, but others as well) are as important as Archaea in eukaryogenesis. Mitochondria are of alphaproteobacterial origin, nuclei of chimeric (archaeal and bacterial), and plastids of cyanobacterial origin. The strongest signals in eukaryotic genomes are, indeed, proteobacterial, archaeal, and cyanobacterial (Pisani et al., 2007; Ku et al., 2015). Because of the combination of archaeal and bacterial features exhibited by eukaryotes, they should not be assigned to a higher taxon along with any of their ancestors.

Eukaryotes exhibit a unique mixture of prokaryotic features, most of which can be traced back to either Archaea or Bacteria. Unlike prokaryotes, eukaryotes do not exchange genes via LGT, but by sexual reproduction (Ku et al., 2015). An archaeon is known to have been the host of the eukaryote-forming endosymbiosis, contributing genetic machinery and ribosomal DNA (Esser et al., 2004; Thiergart et al., 2012; Gould et al., 2016). There is an interesting hypothesis stating that eukaryotic membranes originated from bacterial vesicle secretion (Gould et al., 2016). The genes encoded in the nucleus are as bacterial as they are archaeal. A larger part of the eukaryotic genome has bacterial homologs (Esser et al., 2004; Brueckner and Martin, 2020) that most likely originated from the EGT (endosymbiotic gene transfer) with the proto-mitochondrion ancestor (Brueckner and Martin, 2020), whereas archaeal genes are less numerous in eukaryotic genome, but also important (Pisani et al., 2007; Brueckner and Martin, 2020). The origin of mitochondrion was a prerequisite for the existence of sexual reproduction and meiosis. These processes required large amounts of energy (ATP), and no known prokaryotic cell is able to produce such amount of ATP (Garg and Martin, 2016). Some authors still dispute the uniqueness of eukaryogenesis and the importance of mitochondria in the definition of eukaryotes (e.g., Booth and Doolittle, 2015; Lynch and Marinov, 2016).

We think that the bacterial contribution to eukaryogenesis should not be neglected in view of the facts that: (1) mitochondria, whose presence is a eukaryotic synapomorphy, represents the true descendant of Alphaproteobacteria, (2) most of the eukaryotic nuclear DNA originated via gene transfer from bacteria, and (3) all eukaryotic membranes may be of bacterial origin.

CONCLUDING THOUGHTS

Because of the polyphyletic origin of the eukaryotic monophylum, eukaryogenesis within prokaryotes is not comparable with mammal origin within paraphyletic reptiles. Both synapomorphies and plesiomorphies represent apomorphies and are indeed suitable for defining monophyletic (holophyletic and paraphyletic) groups. Alphaproteobacteria (Bacteria) and Asgard (Archaea) are the ancestors of LECA (the Last Eukaryotic Common Ancestor). The presence of ESPs in Asgard does not dispute the polyphyletic origin of eukaryotes; it only further corroborates it. "Candidatus *Prometheoarchaeum syntrophicum*" is the closest relative to eukaryotes and the only Asgard with available microscopy data. This newly discovered species has a prokaryotic cell organization and does not exhibit features of eukaryotic complexity (nucleus, mitochondrion, meiotic cycle), and thus, it does not belong to Eukaryomorpha.

Along with Cyanobacteria, non-photosynthetic eukaryotes are the ancestors of the primary photosynthetic eukaryotes (archaeplastidians). Non-photosynthetic eukaryotes are not the ancestors of plastids, hence LECA is not the only ancestor of the extant eukaryotic diversity. Eukaryotes are monophyletic by definition, as they have a single ancestor, LECA. They are also holophyletic as all LECA's descendants belong to the same group. They are polyphyletic as well since they exhibit numerous symbioses and anastomoses in the tree of life.

Symbiogenesis will always be one of the major forces driving eukaryotic evolution. A group of polyphyletic origin, such as eukaryotes, should not be assigned to a higher taxon that contains its single parent, as is the case with Eukaryomorpha.

AUTHOR CONTRIBUTIONS

JS contributed to conceptualization, investigation, data curation, writing (original draft), and visualization. DF was responsible for the validation, resources, writing (review and editing), and supervision. All authors contributed to the article and approved the submitted version.

FUNDING

This paper was funded by Volkswagen Stiftung Grant (grant number 93 046) to William F. Martin.

ACKNOWLEDGMENTS

Thanks to Sriram Garg for bringing our attention to Eukaryomorpha as a category. We are thankful to William Martin and Sven Gould for their comments and critiques which improved the quality of this manuscript. We also thank Dora Papković for proofreading the manuscript.

REFERENCES

Adl, S. M., Bass, D., Lane, C. E., Lukeš, J., Schoch, C. L., Smirnov, A., et al. (2019). Revisions to the classification, nomenclature, and diversity of eukaryotes. *J. Eukaryot. Microbiol.* 66, 4–119. doi: 10.1111/jeu.12691

Akil, C., and Robinson, R. C. (2018). Genomes of Asgard archaea encode profilins that regulate actin. *Nature* 562, 439–443. doi: 10.1038/s41586-018-0548-6

Akil, C., Tran, L. T., Orhant-Prioux, M., Baskaran, Y., Manser, E., Blanchoin, L., et al. (2019). Complex eukaryotic-like actin regulation systems from Asgard archaea. *bioRxiv [Preprint]*. 768580. doi: 10.1101/768580

Ashlock, P. D. (1971). Monophyly and associated terms. *Syst. Zool.* 20, 63–69. doi: 10.2307/2412223

Ashlock, P. D. (1972). Monophyly again. *Syst. Zool.* 21, 430–438. doi: 10.2307/2412435

Ashlock, P. D. (1974). The uses of cladistics. *Annu. Rev. Ecol. Evol. S* 5, 81–89. doi: 10.1146/annurev.es.05.110174.000501

Ashlock, P. D. (1979). An evolutionary systematist's view of classification. *Syst. Zool.* 28, 441–450. doi: 10.2307/2412559

Bonfante, P., and Desirò, A. (2017). Who lives in a fungus? The diversity, origins and functions of fungal endobacteria living in *Mucoromycota*. *ISME J.* 11:1727. doi: 10.1038/ismej.2017.21

Booth, A., and Doolittle, W. F. (2015). Eukaryogenesis, how special really?. *Proc. Natl. Acad. Sci. U.S.A.* 112, 10278–10285. doi: 10.1073/pnas.1421376112

Brueckner, J., and Martin, W. F. (2020). Bacterial genes outnumber archaeal genes in eukaryotic genomes. *Genome Biol. Evol.* 12, 282–292. doi: 10.1093/gbe/evaa047

Costello, M., May, R., and Stork, N. (2013). Can we name Earth's species before they go extinct? *Science* 339, 413–416. doi: 10.1126/science.1230318

de Bary, A. (1879). *Die Erscheinung der Symbiose*. Strassburg: Verlag von Karl J. Trubner.

Doolittle, W. F. (2014). How natural a kind is "eukaryote?". *Cold Spring Harbor Perspect. Biol.* 6:a015974. doi: 10.1101/cshperspect.a015974

Doolittle, W. F. (2020). Evolution: two domains of life or three?. *Curr. Biol.* 30, R177–R179. doi: 10.1016/j.cub.2020.01.010

Envall, M. (2008). On the difference between mono-, holo-, and paraphyletic groups: a consistent distinction of process and pattern. *Biol. J. Linn. Soc.* 94, 217–220. doi: 10.1111/j.1095-8312.2008.00984.x

Esser, C., Ahmadinejad, N., Wiegand, C., Rotte, C., Sebastiani, F., Gelius-Dietrich, G., et al. (2004). A genome phylogeny for mitochondria among α-proteobacteria and a predominantly eubacterial ancestry of yeast nuclear genes. *Mol. Biol. Evol.* 21, 1643–1660. doi: 10.1093/molbev/msh160

Fournier, G. P., and Poole, A. M. (2018). A briefly argued case that Asgard archaea are part of the eukaryote tree. *Front. Microbiol.* 9:1896. doi: 10.3389/fmicb.2018.01896

Garg, S., and Martin, W. F. (2016). Mitochondria, the cell cycle, and the origin of sex via a syncytial eukaryote common ancestor. *Genome Biol. Evol.* 8, 1950–1970. doi: 10.1093/gbe/evw136

Gottsberger, B. (2007). *Interspecific Hybridization Between the Grasshoppers Chorthippus biguttulus and Chorthippus Brunneus (Acrididae; Gomphocerinae)* (PhD thesis). Friedrich-Alexander-Universität Erlangen-Nürnberg.

Gould, S. B., Garg, S. G., and Martin, W. F. (2016). Bacterial vesicle secretion and the evolutionary origin of the eukaryotic endomembrane system. *Trends Microbiol.* 24, 525–534. doi: 10.1016/j.tim.2016.03.005

Grant, P. R., and Grant, B. R. (2008). *How and why Species Multiply: The Radiation of Darwin's Finches*. New Jersey, NJ: Princeton University Press.

Hennig, W. (1950). *Grundzüge Einer Theorie der Phylogenetischen Systematik*. Berlin: Deutcher Zentralverlag.

Hennig, W. (1966). *Phylogenetic Systematics*. Urbana: University of Illinois Press.

Imachi, H., Nobu, M. K., Nakahara, N., Morono, Y., Ogawara, M., Takaki, Y., et al. (2020). Isolation of an archaeon at the prokaryote-eukaryote interface. *Nature* 577, 519–525. doi: 10.1038/s41586-019-1916-6

Koonin, E. V. (2010). The origin and early evolution of eukaryotes in the light of phylogenomics. *Genome Biol.* 11:209. doi: 10.1186/gb-2010-11-5-209

Koumandou, V. L., Wickstead, B., Ginger, M. L., Van Der Giezen, M., Dacks, J. B., and Field, M. C. (2013). Molecular paleontology and complexity in the last eukaryotic common ancestor. *Crit. Rev. Biochem. Mol. Biol.* 48, 373–396. doi: 10.3109/10409238.2013.821444

Ku, C., Nelson-Sathi, S., Roettger, M., Sousa, F. L., Lockhart, P. J., Bryant, D., et al. (2015). Endosymbiotic origin and differential loss of eukaryotic genes. *Nature* 524, 427–432. doi: 10.1038/nature14963

Lindsay, W. L. (1856). *A Popular History of British Lichens, Comprising an Account of Their Structure, Reproduction, Uses, Distribution, and Classification*. London: Lovell Reeve and Co.

López, D., Vlamakis, H., and Kolter, R. (2010). Biofilms. *Cold. Spring. Harb. Perspect. Biol.* 2:a000398. doi: 10.1101/cshperspect.a000398

Lutzoni, F., and Miadlikowska, J. (2009). Lichens. *Curr. Biol.* 19, R502–R503. doi: 10.1016/j.cub.2009.04.034

Lynch, M., and Marinov, G. K. (2016). Reply to Lane and Martin: Mitochondria do not boost the bioenergetic capacity of eukaryotic cells. *Proc. Natl. Acad. Sci. U.S.A.* 113, E667–E668. doi: 10.1073/pnas.1523394113

MacLeod, F., Kindler, G. S., Wong, H. L., Chen, R., and Burns, B. P. (2019). Asgard archaea: Diversity, function, and evolutionary implications in a range of microbiomes. *AIMS Microbiol.* 5, 48–61. doi: 10.3934/microbiol.2019.1.48

Margulis, L. (1991). *Symbiosis as a Source of Evolutionary Innovation. Speciation and Morphogenesis*. Cambridge MIT Press.

Martin, W. F. (1999). A briefly argued case that mitochondria and plastids are descendants of endosymbionts, but that the nuclear compartment is not. *Proc. R. Soc. B.* 266, 1387–1395. doi: 10.1098/rspb.1999.0792

McFadden, G. I. (2001). Primary and secondary endosymbiosis and the origin of plastids. *J. Phycol.* 37, 951–959. doi: 10.1046/j.1529-8817.2001.01126.x

McInerney, J. O., O'Connell, M., and Pisani, D. (2014). The hybrid nature of the eukaryota and a consilient view of life on Earth. *Nat. Rev. Microbiol.* 12, 449–455. doi: 10.1038/nrmicro3271

Méheust, R., Watson, A. K., Lapointe, F.-J., Papke, R. T., Lopez, P., and Bapteste, E. (2018). Hundreds of novel composite genes and chimeric genes with bacterial origins contributed to haloarchaeal evolution. *Genom. Biol.* 19:75. doi: 10.1186/s13059-018-1454-9

Meier, J. I., Marques, D. A., Mwaiko, S., Wagner, C. E., Excoffier, L., and Seehausen, O. (2017). Ancient hybridization fuels rapid cichlid fish adaptive radiations. *Nature Comm.* 8:14363. doi: 10.1038/ncomms14363

Mereschkowsky, C. (1905). Über natur und ursprung der chromatophoren im pflanzenreiche. *Biol. Centralbl.* 25, 593–604. doi: 10.1017/S0967026299002231

Muñoz-Gómez, S. A., Wideman, J. G., Roger, A. J., and Slamovits, C. H. (2017). The origin of mitochondrial cristae from alphaproteobacteria. *Mol. Biol. Evol.* 34, 943–956. doi: 10.1093/molbev/msw298

Naumann, M., Schüßler, A., and Bonfante, P. (2010). The obligate endobacteria of arbuscular mycorrhizal fungi are ancient heritable components related to the Mollicutes. *ISME J.* 4:862. doi: 10.1038/ismej.2010.21

Nelson-Sathi, S., Sousa, F. L., Roettger, M., Lozada-Chávez, N., Thiergart, T., Janssen, A., et al. (2015). Origins of major archaeal clades correspond to gene acquisitions from bacteria. *Nature* 517, 77. doi: 10.1038/nature13805

Nielsen, C. (2012). *Animal Evolution: Interrelationships of the Living Phyla*. Oxford: Oxford University Press.

Nowack, E. C., and Melkonian, M. (2010). Endosymbiotic associations within protists. *Phil. Trans. R. Soc. B.* 365, 699–712. doi: 10.1098/rstb.2009.0188

Ozkan, H., Levy, A. A., and Feldman, M. (2001). Allopolyploidy-induced rapid genome evolution in the wheat (*Aegilops–Triticum*) group. *Plant Cell.* 13, 1735–1747. doi: 10.1105/TPC.010082

Pisani, D., Cotton, J. A., and McInerney, J. O. (2007). Supertrees disentangle the chimerical origin of eukaryotic genomes. *Mol. Biol. Evol.* 24, 1752–1760. doi: 10.1093/molbev/msm095

Pough, F. H., Janis, C. M., and Heiser, J. B. (1999). *Vertebrate Life, 5th Edition*. New York, NY: Upper Saddle River, Prentice Hall.

Recuero, E., Buckley, D., García-París, M., Arntzen, J. W., Cogălniceanu, D., and Martínez-Solano, I. (2014). Evolutionary history of *Ichthyosaura alpestris* (Caudata, Salamandridae) inferred from the combined analysis of nuclear and mitochondrial markers. *Mol. Phylogenet. Evol.* 81, 207–220. doi: 10.1016/j.ympev.2014.09.014

Roger, A. J., Muñoz-Gómez, S. A., and Kamikawa, R. (2017). The origin and diversification of mitochondria. *Curr. Biol.* 27, R1177–R1192. doi: 10.1016/j.cub.2017.09.015

Sankararaman, S., Mallick, S., Patterson, N., and Reich, D. (2016). The combined landscape of Denisovan and Neanderthal ancestry in present-day humans. *Curr. Biol.* 26, 1241–1247. doi: 10.1016/j.cub.2016.03.037

Seehausen, O. (2004). Hybridization and adaptive radiation. *Trends Ecol. Evol.* 19, 198–207. doi: 10.1016/j.tree.2004.01.003

Seitz, K. W., Lazar, C. S., Hinrichs, K. U., Teske, A. P., and Baker, B. J. (2016). Genomic reconstruction of a novel, deeply branched sediment archaeal phylum with pathways for acetogenesis and sulfur reduction. *ISME J.* 10, 1696–1705. doi: 10.1038/ismej.2015.233

Spang, A., Saw, J. H., Jørgensen, S. L., Zaremba-Niedzwiedzka, K., Martijn, J., Lind, A. E., et al. (2015). Complex archaea that bridge the gap between prokaryotes and eukaryotes. *Nature* 521, 173. doi: 10.1038/nature14447

Speijer, D., Luke,š, J., and Eli,áš, M. (2015). Sex is a ubiquitous, ancient, and inherent attribute of eukaryotic life. *Proc. Natl. Acad. Sci. U.S.A.* 112, 8827–8834. doi: 10.1073/pnas.1501725112

Spribille, T., Tuovinen, V., Resl, P., Vanderpool, D., Wolinski, H., Aime, M. C., et al. (2016). Basidiomycete yeasts in the cortex of ascomycete macrolichens. *Science* 353, 488–492. doi: 10.1126/science.aaf8287

Thiergart, T., Landan, G., Schenk, M., Dagan, T., and Martin, W. F. (2012). An evolutionary network of genes present in the eukaryote common ancestor polls genomes on eukaryotic and mitochondrial origin. *Genome Biol. Evol.* 4, 466–485. doi: 10.1093/gbe/evs018

Tuovinen, V., Ekman, S., Thor, G., Vanderpool, D., Spribille, T., et al. (2019). Two Basidiomycete fungi in the cortex of wolf lichens. *Curr. Biol.* 29, 476–483. doi: 10.1016/j.cub.2018.12.022

Vogels, G. D., Hoppe, W. F., and Stumm, C. K. (1980). Association of methanogenic bacteria with rumen ciliates. *Appl. Environ. Microbiol.* 40, 608–612. doi: 10.1128/AEM.40.3.608-612.1980

Wake, D. B., Wake, M. H., and Specht, C. D. (2011). Homoplasy: from detecting pattern to determining process and mechanism of evolution. *Science* 331, 1032–1035. doi: 10.1126/science.1188545

Wallin, I. E. (1927). *Symbionticism and the Origin of Species*. Moscow: Ripol Klassik.

Williams, T. A., Foster, P. G., Cox, C. J., and Embley, T. M. (2013). An archaeal origin of eukaryotes supports only two primary domains of life. *Nature* 504, 231–236. doi: 10.1038/nature12779

Woese, C. R., Kandler, O., and Wheelis, M. L. (1990). Towards a natural system of organisms: proposal for the domains *Archaea*, Bacteria, and *Eucarya*. *Proc. Natl. Acad. Sci. U.S.A.* 87, 4576–4579. doi: 10.1073/pnas.87.12.4576

Zaremba-Niedzwiedzka, K., Caceres, E. F., Saw, J. H., Bäckström, D., Juzokaite, L., Vancaester, E., et al. (2017). Asgard archaea illuminate the origin of eukaryotic cellular complexity. *Nature* 541, 353–358. doi: 10.1038/nature21031

Zimorski, V., Ku, C., Martin, W. F., and Gould, S. B. (2014). Endosymbiotic theory for organelle origins. *Curr. Opin. Microbiol.* 22, 38–48. doi: 10.1016/j.mib.2014.09.008

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Gene Duplications Trace Mitochondria to the Onset of Eukaryote Complexity

Fernando D.K. Tria ,*,†,[1] Julia Brueckner,†,[1] Josip Skejo,[1,2] Joana C. Xavier ,[1] Nils Kapust ,[1] Michael Knopp,[1] Jessica L.E. Wimmer,[1] Falk S.P. Nagies,[1] Verena Zimorski,[1] Sven B. Gould,[1] Sriram G. Garg,[1] and William F. Martin[1]

[1]Institute for Molecular Evolution, Heinrich Heine University Düsseldorf, Germany

[2]Faculty of Science, University of Zagreb, Croatia

†These authors contributed equally to this work.

*Corresponding author: E-mail: tria@hhu.de.

Accepted: 14 March 2021

Abstract

The last eukaryote common ancestor (LECA) possessed mitochondria and all key traits that make eukaryotic cells more complex than their prokaryotic ancestors, yet the timing of mitochondrial acquisition and the role of mitochondria in the origin of eukaryote complexity remain debated. Here, we report evidence from gene duplications in LECA indicating an early origin of mitochondria. Among 163,545 duplications in 24,571 gene trees spanning 150 sequenced eukaryotic genomes, we identify 713 gene duplication events that occurred in LECA. LECA's bacterial-derived genes include numerous mitochondrial functions and were duplicated significantly more often than archaeal-derived and eukaryote-specific genes. The surplus of bacterial-derived duplications in LECA most likely reflects the serial copying of genes from the mitochondrial endosymbiont to the archaeal host's chromosomes. Clustering, phylogenies and likelihood ratio tests for 22.4 million genes from 5,655 prokaryotic and 150 eukaryotic genomes reveal no evidence for lineage-specific gene acquisitions in eukaryotes, except from the plastid in the plant lineage. That finding, and the functions of bacterial genes duplicated in LECA, suggests that the bacterial genes in eukaryotes are acquisitions from the mitochondrion, followed by vertical gene evolution and differential loss across eukaryotic lineages, flanked by concomitant lateral gene transfer among prokaryotes. Overall, the data indicate that recurrent gene transfer via the copying of genes from a resident mitochondrial endosymbiont to archaeal host chromosomes preceded the onset of eukaryotic cellular complexity, favoring mitochondria-early over mitochondria-late hypotheses for eukaryote origin.

Key words: evolution, paralogy, gene transfer, endosymbiosis, gene duplication, eukaryote origin.

Significance

The origin of eukaryotes is one of evolution's classic unresolved issues. At the center of debate is the relative timing of two canonical eukaryotic traits: cellular complexity and mitochondria. Gene duplications fostered the evolution of novel eukaryotic traits and serve as a rich phylogenetic resource to address the question. By investigating gene duplications that trace to the last eukaryotic common ancestor we found evidence for mitochondria preceding cellular complexity in eukaryote evolution. Our results demonstrate that gene duplications were already rampant in the last eukaryote common ancestor, and we propose that the vast majority of duplications resulted from cumulative rounds of gene transfers from the mitochondrial ancestor to the genome of the archaeal host cell.

Introduction

The last eukaryote common ancestor (LECA) lived about 1.6 Ba (Betts et al. 2018; Javaux and Lepot 2018). It possessed bacterial lipids, nuclei, sex, an endomembrane system, mitochondria, and all other key traits that make eukaryotic cells more complex than their prokaryotic ancestors (Speijer et al. 2015; Gould et al. 2016; Zachar and Szathmáry 2017; Barlow et al. 2018; Betts et al. 2018). The closest known relatives of the host lineage that acquired the mitochondrion are, however, small obligately symbiotic archaea from enrichment cultures that lack any semblance of eukaryotic cell complexity (Imachi et al. 2020). This steep evolutionary grade separating prokaryotes from eukaryotes increasingly implicates mitochondrial symbiosis at eukaryote origin (Gould et al. 2016; Imachi et al. 2020). Yet despite the availability of thousands of genome sequences, and five decades to ponder Margulis (Margulis et al. 2006) resurrection of endosymbiotic theory (Mereschkowsky 1910; Wallin 1925), the timing, and evolutionary significance of mitochondrial origin remains a polarized debate. Gradualist theories contend that eukaryotes arose from archaea by slow accumulation of eukaryotic traits (Cavalier-Smith 2002; Booth and Doolittle 2015; Hampl et al. 2019) with mitochondria arriving late (Pittis and Gabaldón 2016), whereas symbiotic theories have it that mitochondria initiated the onset of eukaryote complexity in a nonnucleated archaeal host (Imachi et al. 2020) by gene transfers from the organelle (Martin and Müller 1998; Lane and Martin 2010; Gould et al. 2016; Martin et al. 2017).

Information from gene duplications can help to resolve this debate. Gene and genome duplications are a genomic proxy for biological complexity and are the hallmark of eukaryotic genome evolution (Ohno 1970). Gene families that were duplicated during the transition from the first eukaryote common ancestor (FECA) to LECA could potentially shed light on the relative timing of mitochondrial acquisition and eukaryote complexity if they could be inferred in a quantitative rather than piecemeal manner. Duplications of individual gene families (Hittinger and Carroll 2007) and whole genomes (Scannell et al. 2006; Van De Peer et al. 2009) have occurred throughout eukaryote evolution. This is in stark contrast to the situation in prokaryotes, where gene duplications are rare at best (Treangen and Rocha 2011) and whole-genome duplications of the kind found in eukaryotes are altogether unknown. In an earlier study, Makarova et al. (2005) used a liberal criterion and attributed any gene present in two major eukaryotic lineages as present in LECA. Their approach overlooks eukaryotic lineage phylogeny, leading to the inference of 4,137 families that might have been duplicated in LECA. More recently, Vosseberg et al. (2021) examined nodes in trees derived from protein domains that could be scored as duplications among the 7,447–21,840 genes that they estimated to have been present in LECA and used branch lengths to estimate the timing of duplication events. However, they did not report integer numbers for duplications because of their approach based on the analyses of very large protein-domain trees instead of discrete protein-coding gene trees. Here, we addressed the problem of which, what kind of, and how many genes were duplicated in LECA and discuss the implications of our findings for the mitochondria-early versus mitochondria-late debate.

Results and Discussion

To ascertain when the process of gene duplication in eukaryote genome evolution commenced and whether mitochondria might have been involved in that process, we inferred all gene duplications among the 1,848,936 protein-coding genes present in 150 sequenced eukaryotic genomes. For this, we first clustered all eukaryotic proteins using a low stringency clustering threshold of 25% global amino acid identity (see Materials and Methods) in order to recover the full spectrum of eukaryotic gene duplications in both highly conserved and poorly conserved gene families. We emphasize that we employed a clustering threshold of 25% amino acid identity because our procedure was designed to allow for the construction of alignments and phylogenetic trees for each cluster. The 25% threshold keeps the alignments and trees out of the "twilight zone" of sequence identity (Jeffroy et al. 2006), where alignment and phylogeny artifacts based on comparisons of nonhomologous amino acid positions arise.

We then identified all genes that were duplicated across 150 sequenced eukaryotic genomes. In principle, genes present only in one copy in any genome could have also undergone duplication, with losses leading to single-copy status. Quantifying duplications in such cases are extremely topology-dependent. We therefore focused our attention on genes for which topology-independent evidence for duplications existed, that is, genes that were present in more than one copy in at least one genome. Eukaryotic gene duplications were found in all six supergroups: Archaeplastida, Opisthokonta, Mycetozoa, Hacrobia, SAR, and Excavata (Adl et al. 2012), whereby 941,268 of all eukaryotic protein-coding genes, or nearly half the total, exist as multiple copies in at least one genome. These are distributed across 239,012 gene families, which we designate as multicopy gene families. However, 89.7% of these gene families harbor only recent gene duplications, restricted to a single eukaryotic genome (inparalogs). The remaining 24,571 families (10.3%) harbor multiple copies in at least two eukaryotic genomes, with variable distribution across the supergroups (fig. 1). Opisthokonts (animals and fungi) together harbor a total of 22,410 multicopy gene families present in at least two genomes. The animal lineage harbors 19,530 multicopy gene families, the largest number of any lineage sampled, followed by the plant lineage (Archaeplastida) with 6,495 multicopy gene families. Of particular importance for the present study, among the 24,571 multicopy gene families, we

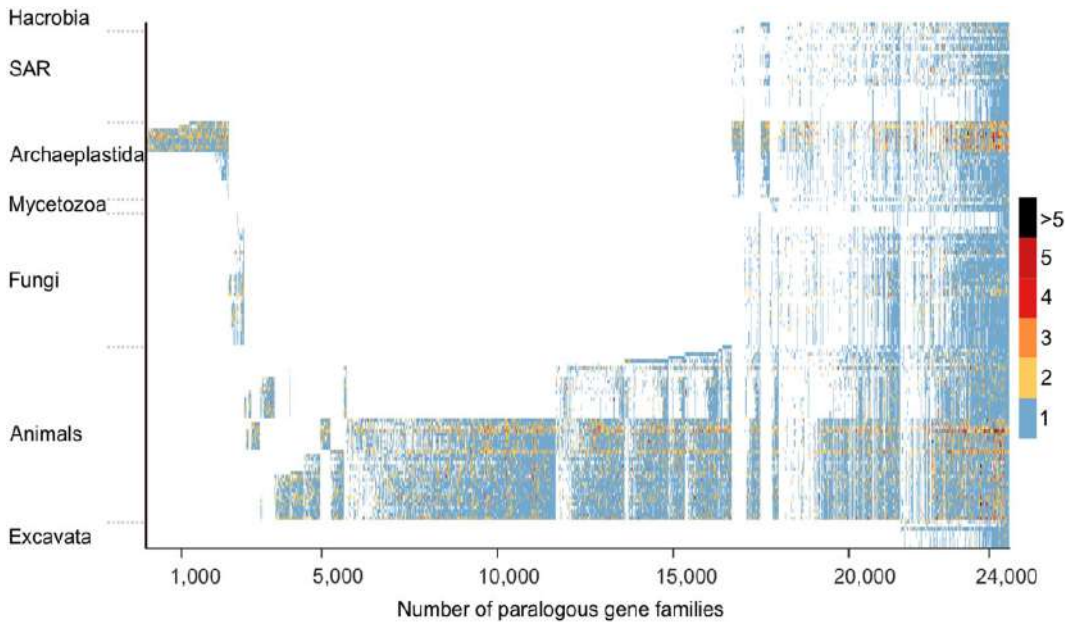


Fig. 1.—Distribution of multicopy genes across 150 eukaryotic genomes. All eukaryotic protein-coding genes were clustered, aligned, and used for phylogenetic inferences. The resulting gene families present as multiple copies in more than one genome are plotted (see Materials and Methods). The figure displays the 24,571 multicopy gene families (horizontal axis) and the colored scale indicates the number of gene copies in each eukaryotic genome (vertical axis). The genomes were sorted according to a reference species tree (supplementary data 7) and taxonomic classifications were taken from NCBI. Animals and fungi together form the opisthokont supergroup.

identified 1,823 that are present as multiple copies in at least one genome from all six supergroups and are thus potential candidates of gene duplications tracing to LECA. In order to distinguish between the possibility of 1) duplications within supergroups after diversification from LECA and 2) duplications giving rise to multiple copies in the genome of LECA, we used phylogenetic trees.

To infer the relative phylogenetic timing of eukaryotic gene duplication events, we focused our attention on the individual protein alignments and maximum-likelihood trees for all 24,571 gene families with paralogs in at least two eukaryotic genomes. We then assigned gene duplications in each tree to the most recent internal node possible, allowing for multiple gene duplication events and losses as needed (see Materials and Methods) and permitting any branching order of supergroups. This approach minimized the number of inferred duplication events and identified a total of 163,545 gene duplications, 160,676 of which generated paralogs within a single supergroup (inparalogs at the supergroup-level). An additional 2,869 gene duplication events trace to the common ancestor of at least two supergroups (fig. 2*a* and supplementary table 1). The most notable result however was the identification of 713 gene duplication events distributed in 475 gene trees that generated paralogs in the genome of LECA before eukaryotic supergroups diverged. For these 475 gene trees, the resulting LECA paralogs are retained in at least one genome from all six supergroups, as indicated in red in figure 2*a*. The sample of 475 genes provides a conservative estimate of genes that duplicated in LECA. Among the 1,823 gene families having multiple copies in members of all six supergroups, note that only in 475 families (26%) do the duplications actually trace to LECA in the trees. These results indicate that most duplications in eukaryotes are lineage specific (figs. 1 and 2), and furthermore raise caveats regarding earlier estimates of duplications in LECA (Makarova et al. 2005; Vosseberg et al. 2021) based on more permissive criteria.

LECA's Duplications Constrain the Position of the Eukaryotic Root

The six supergroups plus LECA at the root represent a seven-taxon tree with the terminal edges bearing 97% of gene duplication events (fig. 2). Gene duplications that map to internal branches of the rooted supergroup tree can result from duplications in LECA followed by vertical inheritance and differential loss in some supergroups, or they result from more recent duplications following the divergence from LECA. Branches that explain the most duplications are likely to reflect the natural supergroup phylogeny, because support for conflicting branches is generated by random nonphylogenetic patterns of independent gene losses (Van De Peer et al. 2009). There is a strong phylogenetic signal contained within the eukaryotic gene duplication data (fig. 2). Among all possible internal branches, those supported by the most frequent

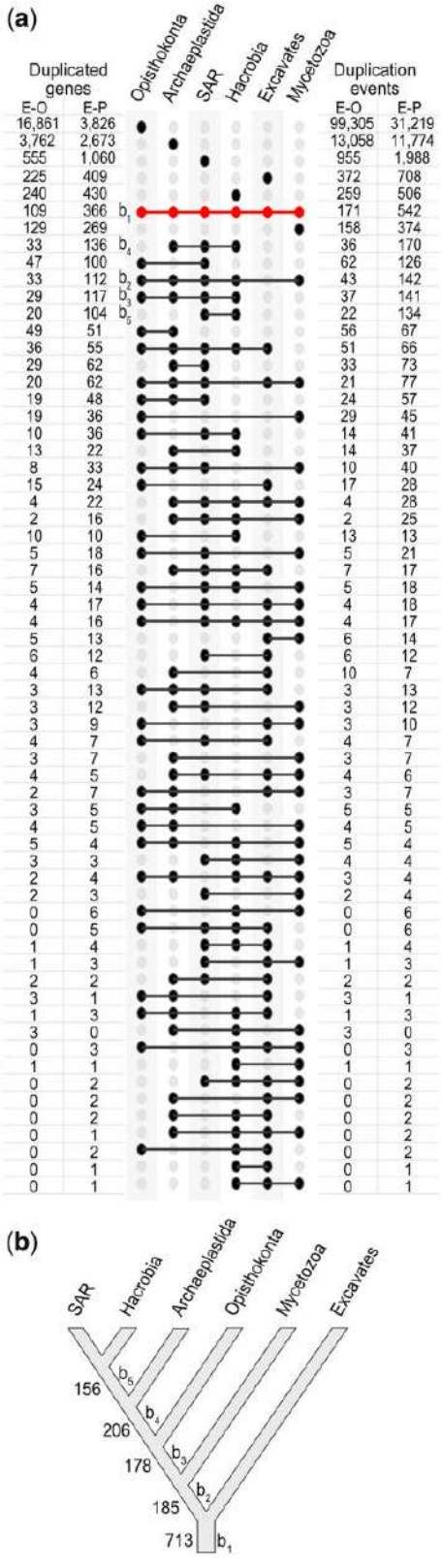


Fig. 2.—Distribution of paralogs descending from gene duplications across six eukaryotic supergroups. (*a*) The figure shows the distribution of paralogs resulting from gene duplications in eukaryotic-specific genes (E-O) and eukaryotic genes with prokaryotic homologs (E-P) (see Materials and Methods for details). Duplicated genes refer to the numbers of gene trees with at least one duplication event with descendant paralogs across the supergroups (filled circles in the center). Number of duplication events refers to the total number of gene duplications. The red row circles indicate gene duplications with descendant paralogs in species from all six supergroups and, thus, tracing to LECA regardless of the eukaryotic phylogeny. An early study assigned 4,137 duplicated gene families to LECA but attributed all copies present in any two major eukaryotic groups to LECA (Makarova et al. 2005). In the present sample, we find 2,869 gene duplication events that trace to the common ancestor of at least two supergroups. Our stringent criterion requiring paralog presence in all six supergroups leaves 713 duplications in 475 gene families in LECA. (*b*) Rooted phylogeny of eukaryotic supergroups that maximizes compatibility with gene duplications. Gene duplications mapping to five edges are shown (b_1, b_2, . . . , b_5). The tree represents almost exactly all edges containing the most duplications, the exception is the branch joining Hacrobia and SAR because the alternative branch joining SAR and Opisthokonta is better supported. However, the resulting subtree ((Opisthokonta, SAR),(Archaeplastida, Hacrobia)) accounts for 249 duplications, fewer than the (Opisthokonta,(Archaeplastida,(SAR, Hacrobia))) subtree shown (262 duplications). The position of the root identifies additional gene duplications tracing to LECA (table 1 and supplementary table 4).

duplications are compatible with the tree in figure 2*b*, which places the eukaryotic root on the branch separating Excavates from other supergroups, as implicated in previous studies of concatenated protein sequences (Hampl et al. 2009; He et al. 2014). However, massive gene loss in specific supergroups (in excavates, e.g., see fig. 1) could impair identification of the eukaryotic root (Zmasek and Godzik 2011; Ku et al. 2015; Albalat and Cañestro 2016). Indeed, the high frequency of duplications that trace to LECA readily explains why resolution of deep eukaryotic phylogeny or the position of the eukaryotic root with traditional phylogenomic approaches (Ren et al. 2016) is so difficult (see also supplementary table 2): LECA was replete with duplications and paralogy. Paralogy imposes conflicting signals onto phylogenetic systematics, but gene duplications harbor novel phylogenetic information in their own right (fig. 2), as shared gene duplications discriminate between alternative eukaryote supergroup relationships.

Eukaryotic Duplications Are Not Transferred across Supergroups

Like the nucleus, mitochondria, and other eukaryotic traits (Speijer et al. 2015; Gould et al. 2016; Zachar and Szathmáry 2017; Barlow et al. 2018; Betts et al. 2018; Imachi et al. 2020), the lineage-specific accrual of gene and genome duplications distinguish eukaryotes from prokaryotes (Ohno 1917; Scannell 2006; Hittinger and Carroll 2007; Van De Peer et al. 2009; Treangen and Rocha 2011). Nonetheless, one might argue that the distribution of duplications observed here does not reflect lineage-dependent processes at all, but lateral gene transfers (LGTs) among eukaryotes instead

(Andersson et al. 2003; Keeling and Palmer 2008; Leger et al. 2018). That is, a duplication could, in theory, originate in one supergroup and one or more gene copies could subsequently be distributed among other supergroups via eukaryote-to-eukaryote LGT. However, were that theoretical possibility true then neither duplications, nor any trait, nor any gene could be traced to LECA because all traits and genes in eukaryotes could, in the extreme, simply reflect 1.6 Byr of lineage-specific invention within one supergroup followed by lateral gene traffic among eukaryotes rather than descent with modification (Andersson et al. 2003; Keeling and Palmer 2008; Leger et al. 2018).

However, the present data themselves exclude the deeply improbable eukaryote-to-eukaryote lateral duplication transfer theory in a subtle but strikingly clear manner. How so? Figures 1 and 2*a* show that 30,439 gene lineages bearing duplications (93% of the total) are restricted in their distribution to "only one supergroup," whereas only 2,245 (7% of the total) are shared among two to five supergroups. That is, only 7% of the duplications are shared across supergroups, hence they are the only possible candidates for LGT among supergroups. For the sake of argument, let us entertain the extreme assumption that *all* 2,245 patterns of shared but nonuniversal duplications involved intersupergroup LGT, recalling that there is no intersupergroup LGT in 93% of the genes (fig. 2 and supplementary table 1). With that generous assumption, the intersupergroup LGT frequency would be maximally 7%. That is an extreme upper bound, though, because the observed 93% frequency for duplicates that are supergroup specific and thus have absolutely no observable intersupergroup LGT should apply equally to the 7% of duplications shared across supergroups. Thus, the more realistic maximum estimate is that 0.49% of duplications (7% of 7%) might have been generated by intersupergroup LGT. This estimate is based solely upon the distribution of the duplicates and the premise that eukaryote supergroups are monophyletic. As it concerns the 475 genes with duplications that trace to LECA (fig. 2 and supplementary table 1), this means that 0.49% out of 475, or about 2.3 genes in our data might have been caused by intersupergroup LGT. That is a very low frequency and is consistent with independent genome-wide phylogenetic tests presented previously (Ku et al. 2015) for the paucity of eukaryote-to-eukaryote LGT. If we count duplication events (fig. 2*a*, right panel) rather than gene lineages (fig. 2*a*, left panel), the picture is even more vertical, because 98% of the events are supergroup-specific, hence lacking any patterns that could reflect LGT, meaning that maximally 0.04% (2% of 2%) or 0.19 duplications among 475 (which rounds to zero genes) could be the result of lateral transfer. The supergroup-specific distributions of duplications themselves thus provide very strong evidence that the distribution of duplicated genes in eukaryotes is not the result of eukaryote-to-eukaryote LGT phenomena (Andersson et al. 2003; Keeling and Palmer 2008; Leger et al. 2018) but the result of vertical evolution within supergroups accompanied by gene birth, death (Nei et al. 1997), and differential gene loss (Ku et al. 2015).

LECA's Duplications Support an Early Mitochondrion

Arguably, the timing of mitochondrial origin is the central so far unresolved issue at the heart of eukaryote origin. Several alternative theories for eukaryogenesis have been proposed (reviewed in Martin et al. 2001; Embley and Martin 2006; Poole and Gribaldo 2014; López-García and Moreira 2015; Eme 2017). Symbiogenic theories posit a causal role for mitochondrial endosymbiosis at the origin of cellular eukaryotic complexity (Lane and Martin 2010) with the host being a garden variety archaeon (Martin and Müller 1998). Gradualist theories posit an autogenous origin of eukaryote cell complexity with little or no contribution of the mitochondrion to eukaryogenesis (Cavalier-Smith 2002; Gray 2014). Intermediate theories posit the existence of endosymbioses prior to the origin of mitochondria. These include an endosymbiotic origin of the nucleus (Lake and Rivera 1994), an endosymbiotic origin of peroxisomes (de Duve 2007), an endosymbiotic origin of flagella (Margulis et al. 2000), the lateral acquisition of the cytoskeleton (Doolittle 1998) or, more liberally, additional symbioses preceding the mitochondrion in unconstrained numbers, as long as each symbiosis "explains the origin of any eukaryotic innovation as a response to an endosymbiotic interaction" (Gabaldón 2018). Most current theories posit an origin of the host from archaea (Martin et al. 2015; Spang et al. 2015; Zaremba-Niedzwiedzka et al. 2017; Imachi 2020), though theories for eukaryote origins from actinobacteria (Cavalier-Smith 2002), and planctomycetes (Cavalier-Smith and Chao 2020) are discussed. Notwithstanding such diversity of views, the main divide among theories for eukaryote origin remains the relative timing of mitochondrial origin, that is did the mitochondrion initiate or culminate eukaryote origin (Martin et al. 2001; Embley and Martin 2006; Poole and Gribaldo 2014; López-García and Moreira 2015; Eme et al. 2017)? Alternative theories for eukaryote origin generate distinct predictions about the nature of gene duplications in LECA.

Gradualist theories entailing an archaeal host (Cavalier-Smith 2002; Booth and Doolittle 2015; Pittis and Gabaldón 2016; Hampl et al. 2019) predict genes of archaeal origin and eukaryote-specific genes to have undergone numerous duplications during the origin of eukaryote complexity, prior to the acquisition of the mitochondrion. In that case, the mitochondrion arose late, hence bacterial-derived genes would have accumulated fewer duplications in LECA than archaeal-derived or eukaryote-specific genes (fig. 3*a*). Models invoking gradual lateral gene transfers (LGT) from ingested (phagocytosed) food prokaryotes prior to the origin of mitochondria (Doolittle 1998) also predict more duplications in archaeal-derived and eukaryote-specific genes to underpin the origin

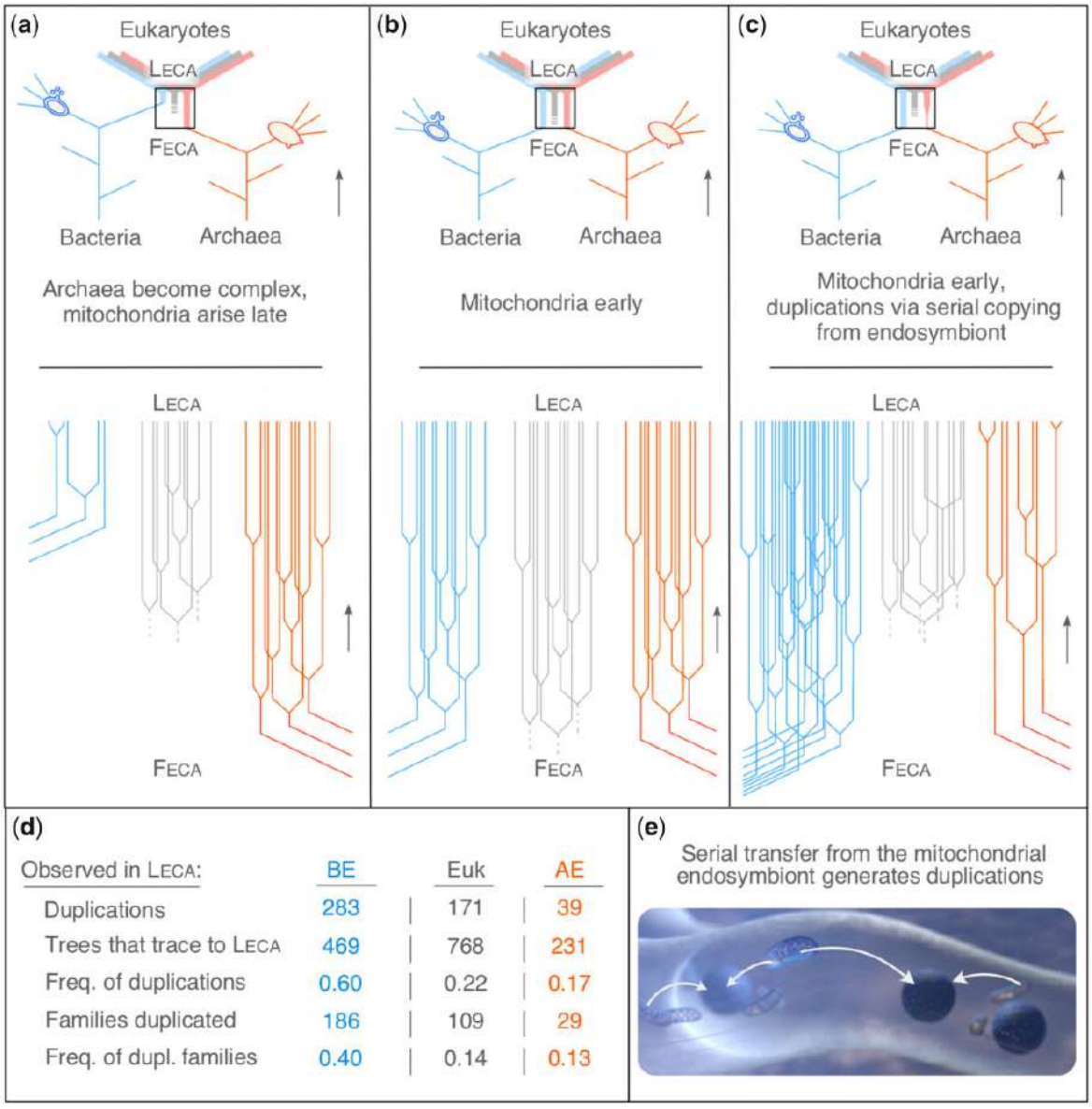


FIG. 3.—Alternative models for eukaryote origin generate different predictions with respect to duplications. In each panel, gene duplications during the FECA to LECA transition (boxed in upper portion) are enlarged in the lower portion of the panel. (*a*) Cellular complexity and genome expansion in an archaeal host predate the origin of mitochondria. (*b*) Mitochondria enter the eukaryotic lineage early, duplications in mitochondrial-derived, host-derived, and eukaryotic-specific genes occur, genome expansion affects all genes equally. (*c*) Gene transfers from a resident endosymbiont generate duplications in genes of bacterial origin in an archaeal host. (*d*) Observed frequencies from gene duplications that trace to LECA (see supplementary table 1). BE refers to eukaryotic genes with bacterial homologs only; AE refers to eukaryotic genes with archaeal homologs only; and Euk refers to eukaryotic genes without prokaryotic homologs. (*e*) Schematic representation of serial gene transfers from the mitochondrion (white arrows) to the host's chromosomes.

of phagocytotic feeding, but do not predict duplications specifically among acquired genes (whether from bacterial or archaeal food) because each ingestion contributes genes only once.

By contrast, transfers from the endosymbiotic ancestors of organelles continuously generated gene duplications in the host's chromosomes (Timmis et al. 2004; Allen 2015), a process that continues to the present day in eukaryotic genomes (Timmis et al. 2004; Portugez et al. 2018). Symbiogenic theories posit that the host that acquired the mitochondrion was an archaeon of normal prokaryotic complexity (Martin and Müller 1998; Lane and Martin 2010; Gould et al. 2016; Martin et al. 2017; Imachi et al. 2020) and hence lacked duplications underpinning eukaryote complexity. There are examples known in which bacteria grow in intimate association with archaea (Imachi et al. 2020) and in which

prokaryotes become endosymbionts within other prokaryotic cells (Martin et al. 2017). However, there are two different ways in which mitochondria could promote the accumulation of duplications. If energetic constraints (Lane and Martin 2010) were the sole factor permitting genome expansion, duplications would accrue in all genes regardless of their origin, such that gene duplications in the wake of mitochondrial origin should be equally common in genes of bacterial, archaeal, or eukaryote-specific origin, respectively (fig. 3*b*). If, on the other hand, the role of mitochondria in gene duplications was mechanistic rather than purely energetic, genes of mitochondrial origin should preferentially undergo duplication. This is because the mechanism of gene transfers from resident organelles involve endosymbiont lysis and the "copying" (Allen 2015) of organelle genomes to the host's chromosomes followed by recombination and mutation (Portugez et al. 2018). Gene transfers from resident endosymbionts specifically generate duplications of endosymbiont genes because new copies of the same genes are recurrently transferred (Timmis et al. 2004; Allen 2015) (fig. 3*c*).

The duplications in LECA reveal a vast excess of duplications in LECA's bacterial-derived genes relative to archaeal-derived and eukaryote-specific genes (fig. 3*d*). Of all gene families tracing to LECA, 26% experienced at least one duplication event during the transition to LECA from FECA. Notably, the excess proportion of duplicates among genes of bacterial origin is significant as judged by the two-tailed binomial test ($P=1.3\times10^{-10}$; proportion of duplicates at 95% CI=[35–44%]; df=1). On the other hand, genes of archaeal origin show significantly fewer duplicates ($P=8.4\times10^{-7}$; proportion of duplicates 95% CI=[8–17%]; df=1) with the proportion of duplicates being similar to eukaryote-specific genes (fig. 3*d*).

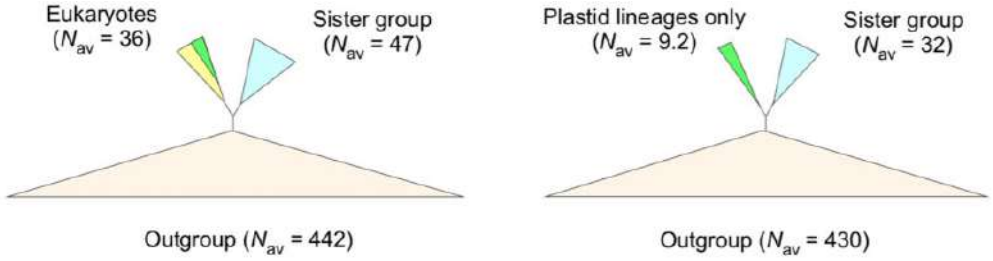
Do Bacterial Genes in LECA Stem from the Mitochondrion?

If bacterial genes in LECA stem from the mitochondrion, as opposed to 1) eukaryote-to-eukaryote gene transfers, which were already excluded for >99% of the families with duplications in this data on the basis of their distributions alone, or 2) multiple lineage-specific acquisitions from bacteria via LGT, then the bacterial genes should trace to the eukaryote common ancestor. That is, the eukaryotes should form a monophyletic clade in gene trees that connect prokaryotic and eukaryotic genes. To test this, we generated clusters, alignments, and trees for genes shared by prokaryotes and eukaryotes from 22,471,723 million genes from 5,655 genomes and including 150 eukaryotes (see Materials and Methods). The results from the 2,575 trees that contained at least five prokaryotic and at least two eukaryotic sequences are summarized in figure 4. As with the duplications themselves, eukaryote gene evolution is again vertical. Out of the 2,575 trees only 475 did not recover eukaryotes as monophyletic. However, none of these 475 trees rejected eukaryote monophyly using the Shimodaira–Hasegawa (SH) test (see Materials and Methods) and only 25 trees (1% of the total) rejected eukaryote monophyly using the Kishino–Hasegawa (KH) test. Applying the approximately unbiased (AU) test, only three trees out of 475 rejected eukaryote monophyly. This traces gene origin of ≥1,649 out of the 2,575 genes shared by prokaryotes and eukaryotes to LECA, and the origin of ≤926 genes to the archaeplastidal ancestor because the latter trees contain only photosynthetic eukaryotic lineages (fig. 4*a*).

The 1,649 trees that trace prokaryotic gene origins to LECA fall into two classes with regard to the sister group of the eukaryotic gene: 966 in which the prokaryotic sister group to eukaryotes contained members of only one phylum (a "pure" sister, S_{pure} in fig. 4, 59% of the trees) and those in which the sister to the eukaryotes contained members of more than one phylum (a "mixed" sister, 41% of the trees). The only way to obtain a mixed sister topology of prokaryotic sequences for a eukaryotic gene is via LGT among prokaryotes (Ku and Martin 2016). If we exclude the reality of LGT among prokaryotes, and interpret mixed sister topologies at face value, they would suggest that eukaryotes arose before the diversification of the diverse prokaryotic phyla present in our sample, which would be incompatible with accounts of eukaryote age (Parfrey et al. 2011; Betts et al. 2018), and would furthermore have LECA arising at different times, depending on the membership in the sister group. LGT among the prokaryotic reference sequences in the mixed sister cases (Ku and Martin 2016; Nagies et al. 2020) is clearly the simpler explanation. The pure sister was bacterial in 49% of the trees and archaeal in only 9.5% of the trees. Only in 115 trees (7.0%) was the bacterial pure sister clade alphaproteobacterial. These 115 trees are readily explained because they stem from the mitochondrion, even though the alphaproteobacterial-derived genes in eukaryotes do not all reside in the "same" alphaproteobacterial genome as previously observed (Ku et al. 2015; Nagies et al. 2020), requiring LGT among alphaproteobacteria, at least, to account for the topology. Yet, the crucial and previously underinvestigated issue concerns the remaining 695 pure sister bacterial origin cases (86%) that trace to LECA but reside in a genome that does not carry an alphaproteobacterial taxon label (fig. 4), as recently set forth in a study that examined the phylogeny of only the more conserved fraction of genes shared by prokaryotes and eukaryotes (Nagies et al. 2020).

There are two general ways to explain the 86% of non-alphaproteobacterial genes that trace to LECA. The first is to take one specific aspect of the trees—namely, the taxon label of the sister group—at face value and interpret the data as evidence for independent individual contributions to eukaryotes (via LGT or via multiple resident symbionts) by all of the bacterial phyla in the sample. At the level of the taxa listed in figure 4, that would mean 26 different bacterial donors to LECA in addition to the alphaproteobacterial contribution,

(a)



(b)

	1649 Trees also including plastid-bearing lineages						926 Trees only including plastid-bearing lineages						
Taxon	P_{mono}	S_{non}	S_{mix}	S_{pure}	$S_{p.avg}$	N_{trees}	P_{mono}	S_{non}	S_{mix}	S_{pure}	$S_{p.avg}$	N_{trees}	ID_{gen}
Bacteria													
Acidithiobacillia	0.93	125	11	1	6	137	0.96	75	6	1	1	82	5
Chlorobi	0.80	125	12	0	-	137	0.83	93	8	1	1	102	12
Chlamydiae	0.72	129	15	10	16.5	154	0.75	59	11	10	23.8	80	113
Deinococcus-Thermus	0.65	269	31	12	4.3	312	0.64	128	18	5	8.4	151	26
Synergistetes	0.64	129	18	1	1	148	0.82	72	4	1	1	77	5
Thermotogae	0.64	156	22	2	1	180	0.68	80	9	4	1.3	93	32
Fusobacteria	0.57	166	21	2	11.5	189	0.73	75	6	1	4	82	20
Negativicutes	0.56	204	18	4	1.5	226	0.60	94	8	3	1.7	105	12
Epsilonproteobacteria	0.56	178	27	7	2.4	212	0.70	104	11	4	1.8	119	253
Cyanobacteria	0.55	312	59	24	21.4	395	0.61	173	34	187	20.9	394	94
Tenericutes	0.53	124	17	6	2.2	147	0.35	53	8	4	5.5	65	141
Verrucomicrobia	0.52	212	48	12	1.4	272	0.54	121	18	4	2	143	10
Planctomycetes	0.51	249	46	16	1.5	311	0.59	118	22	13	1.8	153	8
Acidobacteria	0.51	274	45	10	1.2	329	0.60	130	24	4	1	158	8
Nitrospirae	0.50	142	18	1	1	161	0.40	92	8	0	-	100	8
Chloroflexi	0.41	249	45	9	1.1	303	0.45	138	17	9	3.1	164	26
Spirochaetes	0.39	312	58	12	3.2	382	0.47	144	24	12	6.8	180	66
Aquificae	0.39	121	16	1	3	138	0.39	84	6	2	3	92	16
Erysipelotrichia	0.38	144	16	2	1	162	0.31	68	3	0	-	71	8
Bacilli	0.36	439	63	49	8.1	551	0.35	195	30	15	5.5	240	1012
Deltaproteobacteria	0.33	435	149	62	2.8	646	0.38	209	64	44	2.5	317	74
Bacteroidetes	0.33	434	76	66	3.6	576	0.38	186	34	29	3.2	249	185
Actinobacteria	0.33	550	118	157	5.2	825	0.44	289	42	43	9.1	374	593
Clostridia	0.31	364	71	33	2.2	468	0.35	172	27	13	2.7	212	169
Alphaproteobacteria	0.27	515	117	115	31.9	747	0.31	235	48	72	4.5	355	460
Betaproteobacteria	0.27	503	127	49	4.3	679	0.34	246	55	19	2.5	320	493
Gammaproteobacteria	0.25	560	126	147	40.4	833	0.30	273	67	55	23	395	1565
Other bacteria	-	-	-	-	-	420	-	-	-	-	-	210	29
Archaea													
Methanococcales	0.90	190	9	8	4.5	207	0.82	50	7	3	7.7	60	16
Thermococcales	0.88	231	17	11	13.9	259	0.81	43	3	2	7.5	48	27
Methanobacteriales	0.84	218	17	17	7.9	252	0.77	53	4	5	2.6	62	17
Archaeoglobales	0.84	212	24	7	5.3	243	0.78	55	7	6	4.7	68	8
Sulfolobales	0.84	265	6	6	9.5	277	0.87	49	2	1	5	52	29
Methanomicrobiales	0.80	221	23	15	2.7	259	0.67	59	10	13	3.8	82	9
Methanosarcinales	0.48	260	14	29	5.4	303	0.59	75	5	21	7	101	32
Thermoproteales	0.44	223	49	15	3.9	287	0.41	45	5	4	2	54	10
Thermoplasmatales	0.36	220	37	13	2.5	270	0.57	47	18	4	2.5	69	5
Natrialbales	0.28	271	60	13	2	344	0.33	84	15	8	1.8	107	9
Desulfurococcales	0.25	196	45	7	1.9	248	0.35	44	6	5	1.6	55	11
Haloferacales	0.24	259	55	5	2.2	319	0.22	78	13	7	2	98	7
Halobacteriales	0.16	312	28	10	1.2	350	0.18	97	4	8	1.6	109	16
Other archaea	-	-	-	-	-	389	-	-	-	-	-	134	16

Fig. 4.—Identification of prokaryotic sisters in 2,575 eukaryotic–prokaryotic gene trees. (*a*) The individual trees were rooted on the branch leading to the largest prokaryotic clade deriving the sister group to eukaryotes. The average number of sequences in the eukaryotic clade, sister group, and outgroup are indicated. (*b*) The list of bacterial (top) and archaeal (bottom) phyla occurring in the trees exclusive to plant lineages (right) and all other trees (left). Archaeal and bacterial phyla with less than five representative species in the data set were collapsed into "other archaea" and "other bacterial" groups. P_{mono} refers the proportion of trees with a branch (split) separating the species of the phylum from the others; S_{non} refers to the number of occurrence of the phylum only in the outgroup clade; S_{mix} refers to the number of occurrences of the phylum as a mixed sister (more than one phylum in the clade); S_{pure} refers to the number of occurrences of the phylum as pure sister (as the single phylum); $S_{p.avg}$ shows the average size of the sister group when the phylum occurs as a pure sister clade. N_{trees} show the number of occurrences of the phyla across all trees. ID_{gen} refers to the total number of species in each phylum.

Table 1
Functional Categories of Genes Duplicated in LECAa

Category[b]	(*n*)	Bacterial	Archaeal	Universal	Eukaryotic
Metabolism	(141)	64	2	58	17
Protein modification, folding, degradation	(89)	30	8	30	21
Ubiquitination		3	1	—	9
Proteases		9	1	7	1
Kinase/phosphatase/modification		12	6	19	9
Folding		6	—	4	2
Novel eukaryotic traits	(61)	8	4	12	37
Cell cycle		1	1	2	5
Cytoskeleton		4	—	1	19
Endomembrane (ER; Golgi; vesicles)		2	2	8	10
mRNA splicing		1	1	1	3
Mitochondrion	(47)	29	—	9	9
Carbon metabolism	(37)	26	—	11	—
Glycolysis		10	—	5	—
Reserve polysaccharides, other		16	—	6	—
Cytosolic translation	(36)	15	7	10	4
Nucleic acids	(55)	13	7	15	20
Histones		—	—	2	8
RNA		8	3	6	4
DNA		5	4	7	8
Membranes (excluding endomembrane)	(46)	18	1	12	15
Transporters, plasma associated		8	1	9	14
Lipid synthesis		10	—	3	1
Redox	(15)	11	—	4	—
Hypothetical	(229)	81	9	61	78
Total		295	38	222	201

NOTE.—*n*, number of duplicated genes in the corresponding category.

[a]About 475 genes duplicated in LECA and present in all six supergroups plus 281 genes with duplications tracing to the common ancestors of excavates and other supergroups. The annotation, source (bacterial, archaeal, present in bacteria and archaea, eukaryote specific), and the numbers of duplications for each cluster are given in supplementary tables 3 and 4. All categories listed had representatives on both the 475 and the 281 list except mRNA splicing, present in the 475 list only.

[b]The categories do not strictly adhere to KEGG or gene ontology classifications, instead they were chosen to reflect the processes that took place during the FECA to LECA transition. The largest number of duplications in LECA for any individual gene was 12, a dynein chain known from previous studies to have undergone duplications in the common ancestor of plants animals and fungi (Kollmar 2016).

and donations from 13 different archaeal host taxa. With 39 donor phyla, LECA already looks like a grab bag of genes. At the level of genus, the taxon labels of the trees would mean 794 different bacterial donors to LECA under permissive models (Gabaldón 2018), followed by a particularly ad hoc sudden stop of gene influx to eukaryotes after the FECA to LECA transition, because the eukaryotes are monophyletic in these trees. The suggestion of symbiont acquisition and gene transfers without constraints (Gabaldón 2018) carries a hidden and seldom spelled out corollary (Martin 1999). Namely, it entails the strict condition that all of the nonalphaproteobacterial bacterial genes in question not only resided in the genome of members of the 27 different phylum level bacterial taxa at the time of donation to LECA (fig. 4) but furthermore, and crucially, that those genes evolved "vertically" within the chromosomal confines of those respective phyla during the 1.6 Byr since eukaryotes arose. Such unrestricted donor theories (Gabaldón 2018) assume that the present-day phylum taxon label on the gene accurately identifies the donor phylum at the time of transfer. But that is true "if and only if" the gene has been vertically inherited within that phylum (no interphylum LGT) since its donation to LECA (Martin 1999; Esser et al. 2007).

Such theories of unrestricted LGT to eukaryotes with strictly vertical gene evolution among prokaryotes are unlikely and resoundingly rejected by the data. If we look beyond the mere taxon label of the sister group (fig. 4), we see that the putative 27 bacterial donor lineages themselves do not evolve in a vertical manner. The average level of monophyly for bacterial phyla in the 1,649 trees that trace to LUCA is 47% (P_{mono} in fig. 4). Alphaprotebacteria were monophyletic in only 27% of the trees in which they occurred, as were generalists with large genomes such as betaproteobacteria (27%) and actinobacteria (33%). Specialists like chlorobi or chlamydia with more restricted pangenomes were more monophyletic (80% and 72%, respectively). Halophilic archaea, which are known to have acquired many genes from bacteria (Nelson-Sathi et al. 2012), are the least monophyletic prokaryotes

sampled (halobacteriales, 16%, fig. 4). For the 926 genes that, based on their distribution, trace to the archaeplastidal common ancestor (fig. 4, right panel), the bacterial phyla have a higher proportion of monophyly (P=0.006, V = 67, using two-tailed Wilcoxon signed-rank test) than for those genes that trace to LECA. Plastids are younger than mitochondria, hence the genes from the ancestral plastid genome have had less time to migrate across prokaryotic genomes than genes from the ancestral mitochondrial genome. For the prokaryotic genes and phyla in question, evolution is not a vertical process. The bacterial reference system against which to infer the origin of eukaryotic genes that stem from the mitochondrion (or the plastid) is a system of mosaic (Martin 1999) or fluid (Esser et al. 2007) chromosomes. These findings are fully consistent with a recent larger scale investigation of gene verticality across genomes (Nagies et al. 2020).

If we accept the evidence that LGT in prokaryotes is real and if we accept the evidence that mitochondria were once endosymbiotic bacteria, then the expectation for the phylogeny of a gene that was acquired from the mitochondrion is that it traces to a single origin in LECA, which the genes in this study do, but "not" that it traces to alphaproteobacteria. This is because LGT among prokaryotes preceding and subsequent to the origin of mitochondria generates the illusion of many donors by shuffling the taxon labels attached to genes in mosaic bacterial chromosomes (Martin 1999). Most current studies still equate mitochondrial origin with an alphaproteobacterial sister group relationship (Vosseberg et al. 2021), but if we look at all the data, it is clear that such an interpretation is too strict. For example, Vosseberg et al. (2021) found that about 7% of the eukaryotic protein-domains that they examined branched with alphaproteobacterial homologs. But looking beyond the eukaryotic branch, Nagies et al. (2020) found that only about 35% of alphaproteobacterial genes recover alphaproteobacteria monophyly to begin with, and only 16% of the 220 trees in which alphaproteobacteria appeared as the sole sister of all eukaryotes recovered aphaproteobacteria as monophyletic among prokaryotes. To investigate mitochondrial origin from the standpoint of genes, it is not enough to identify the relationship of eukaryote genes to prokaryotic homologs. One has also to investigate the relationship of prokaryotic homologs to each other, because they are the reference system for comparison.

It is because of LGT among prokaryotes that many different groups are implicated as donors of genes to LECA (fig. 4; see also Nagies et al. 2020). There is no evidence independent of gene phylogenies to suggest or support theories for the participation of spirochaetes (Margulis et al. 2006), actinobacteria (Cavalier-Smith 2002), cyanobacteria (Cavalier-Smith 1975), deltaproteobacteria (López-García and Moreira 1999), planctomycetes (Cavalier-Smith and Chao 2020), or multiple donor lineages (Gabaldón 2018) at eukaryote origin (Embley and Martin 2006). One could of course argue that those conflicting theories for contributions from many different prokaryotic lineages are all simultaneously true, but then theories for eukaryogenesis would no longer be constrained by observations in data, and any assertion about eukaryote origin would be permissible as a line of evidence, an untenable state of affairs. The same sets of considerations apply to the cyanobacterial origin of plastids (fig. 4).

If we let go of the belief that sister group relationships between eukaryotic genes and prokaryotic homologs (fig. 4) identify the prokaryotic lineages that donated genes (Martin 1999; Nagies et al. 2020), and take into account the functions encoded by nuclear genes of bacterial origin that were duplicated in LECA (figs. 2 and 4; table 1), the simplest interpretation of the data in our view is that the bacterial duplicates in LECA were donated by the mitochondrion. Other more complicated interpretations are imaginable, but these interpretations do not simultaneously account for the phylogenetic behavior of the bacterial reference phylogeny set, which we have done here and elsewhere (Nagies et al. 2020). Our data furthermore show that eukaryotic genes are of monophyletic origin. With large genomic samples spanning thousands of reference prokaryotic genomes, eukaryotic gene evolution is clearly vertical, both in terms of lineage-specific distribution of gene duplications (fig. 1) and in terms of likelihood ratio tests (Nagies et al. 2020).

Can Positive Selection Explain Excess Bacterial Duplications?

The vast excess of bacterial duplications (fig. 3) and the phylogenies of 2,575 genes that would address the question of gene origin (fig. 4) speak in favor of bacterial acquisition in LECA from a single-resident endosymbiont, the mitochondrion, prior to the origin of eukaryote complexity. Yet one could still imagine numerous individual gene acquisitions in LECA from different donors with a blanket ad hoc hypothesis of "positive selection" increasing the copy number of bacterial-related functions to account for the excess of bacterial-derived duplications (table 1). However, the selection proposal would not explain the excess of bacterial over archaeal or eukaryote-specific genes with the same functional category, as is widely observed in table 1. That is, selection would have to be invoked as a special plea on a bacterial-gene-for-bacterial-gene basis, requiring yet one additional corollary of positive selection for each duplication. Because we observe over 900,000 duplications in the present data, the selection theory to account for duplications carries a burden of too many corollary assumptions.

On the other hand, it is possible that duplications are fundamentally mechanistic in origin, via chromosome mispairing, translocations, genome duplications, or via duplicative transfers from a resident endosymbiont as we argue in this paper. In a context of mosaic, fluid bacterial genomes (Martin 1999; Esser et al. 2007) permitting LGT among prokaryotes (fig. 4) (Nagies et al. 2020), we would require no corollary

assumptions of ad hoc selection. The mechanism of transfer from the endosymbiont generates the excess of bacterial duplications and does so across all functional categories (table 1).

The Functions of Bacterial Duplicates Polarize Events at LECA's Origin

Gene duplications speak to more than phylogeny. Gene duplications are a standard proxy for the evolution of complexity, as diversification of function and form is canonically underpinned by gene family expansion (Ohno 1970). Accordingly, we observe that the morphologically most complex multicellular eukaryotes—plants, animals, and fungi—harbor the largest numbers of duplications (fig. 1). As outlined above, the simplest interpretation of the present data is that complexity started with the mitochondrion. That is not only true for the present data on duplications, is also true from a purely physiological standpoint (Martin et al. 2017) and a bioenergetic standpoint (Lane and Martin 2010).

The functions of genes that were duplicated in LECA help to polarize events in LECA's evolution. For example, LECA had a mitochondrion. LECA's gene duplications in 47 genes with mitochondrial functions include pyruvate dehydrogenase complex, enzymes of the citric acid cycle, components involved in electron transport, a presequence cleavage protease, the ATP–ADP carrier, and seven members of the eukaryote-specific mitochondrial carrier family that facilitates metabolite exchange between the mitochondrion and the cytosol (table 1 and supplementary tables 3 and 4). A recent study estimated that some genes for mitochondrial function were probably duplicated in LECA, but interpreted the data as evidence for mitochondria-intermediate hypothesis (Vosseberg et al. 2021). The methodology used in Vosseberg et al. has major limitations because: 1) the timing of gene duplications was inferred using an approach that equates branch-lengths from phylogenetic trees to time, which is expected to be valid "only if" the evolutionary rate is constant across genes (substitutions and gene loss, for example); 2) prokaryotic sequences were arbitrarily removed from gene trees, inflating the estimates of duplications in genes of archaeal origin; 3) the use of trees for which the same gene sequence can be represented simultaneously in multiple trees, biasing the estimates of duplications and their origin; and 4) the use of too liberal thresholds for gene clustering which result in aberrantly large gene families (see supplementary fig. 5, Supplementary Material online), a potential source of tree reconstruction errors. By contrast, we do not infer time from branch lengths, we did not remove sequences that did not fit our expectations, and gene membership in our gene families is always unique.

Our findings clearly indicate that canonical energy metabolic functions of mitochondria were established in LECA, underscored by additional functions performed by mitochondria in diverse eukaryotic lineages: ten genes for enzymes of the lipid biosynthetic pathway (typically mitochondrial in eukaryotes; Gould et al. 2016), the entire glycolytic pathway (mitochondrial among marine algae; Río Bártulos et al. 2018), and 11 genes involved in redox balance are found among bacterial duplicates. The largest category of duplications with annotated functions concerns metabolism and biosynthesis (table 1).

Many products of bacterial-derived genes operate in the eukaryotic cytosol (Martin et al. 1993; Esser et al. 2004). This is because at the outset of gene transfer from the endosymbiont, there was no mitochondrial protein import machinery (Martin and Müller 1998; Dolezal et al. 2006), and no nucleus, such that the products of genes transferred from the endosymbiont were active in the compartment where the genes were cotranscriptionally translated (French et al. 2007). Gene transfers in large, genome sized fragments from the endosymbiont, as they occur today (Timmis et al. 2004; Portugez 2018), furthermore, permitted entire pathways to be transferred, because the unit of biochemical selection is the pathway and its product, not the individual enzyme (Martin 2010). In the absence of upstream and downstream intermediates and activities in a pathway, the product of a lone transferred gene is generally useless for the cell, expression of the gene becomes a burden, and the transferred gene cannot be fixed (Martin 2010).

Bacterial-derived duplications are present in functions that underpinned the origin of cell compartmentation in LECA (table 1). LECA possessed an endomembrane system consisting of bacterial lipids, as symbiogenic models predict (Gould et al. 2016). Bacterial duplicates, not archaeal duplicates, dominate lipid synthesis and membrane biogenesis (table 1). Functions of bacterial duplicates are also involved in mRNA splicing, a selective force at the origin of the nucleus (Garg and Martin 2016; Eme et al. 2017). The origin of protein import into mitochondria was essential to mitochondrial origin (Dolezal et al. 2006) and encompasses many bacteria-derived duplicates (table 1). LECA's duplicates of bacterial origin are also involved in the origin of eukaryotic-specific traits, including the cell cycle, the cytoskeleton, endomembrane system, and mRNA splicing (table 1). Eukaryote complexity required intracellular molecular movement in the cytosol, which is realized by motor proteins. The protein with the most duplications found in LECA is a light chain dynein with 12 duplications (supplementary table 3), in agreement with previous studies of dynein evolution that document massive dynein gene duplications early in eukaryote evolution (Kollmar 2016).

Notably, ten of the 20 genes encoding cytoskeletal functions that were duplicated in LECA (supplementary tables 3 and 4) encode dynein or kinesin motor proteins (see also Tromer et al. 2019). The bacterial duplicate contribution vastly outnumbers the archaeal contribution to these categories, which are dominated by eukaryote-specific genes, indicating that eukaryotes not only acquired genes, but they also

invented new ones as well (Lane and Martin 2010). Duplications in LECA depict bacterial carbon and energy metabolism in an archaeal host supported by genes that were recurrently donated by a resident symbiont, in line with the predictions of symbiotic theories for the nature of the first eukaryote (Martin and Müller 1998; Martin et al. 2017; Imachi et al. 2020). The functions of duplications are consistent with the predictions of symbiogenic theories but contrast with gradualist theories positing eukaryote origin from an archaeal lineage that attained eukaryote-like complexity in the absence of the mitochondrial endosymbiont (Cavalier-Smith 2002; Booth and Doolittle 2015; Pittis and Gabaldón 2016; Hampl et al. 2019).

What Does This Say about the Biology of LECA?

Gene transfers from the mitochondrion can generate duplications of bacterial-derived genes. What mechanisms promoted genome-wide gene duplication at the prokaryote–eukaryote transition? Population genetic parameters such as variation in population size (Zachar and Szathmáry 2017) apply to prokaryotes and eukaryotes equally, hence they would not affect gene duplications specifically in eukaryotes, but recombination processes (Garg and Martin 2016) in a nucleated cell could. Because LECA possessed meiotic recombination (Speijer et al. 2015), it was able to fuse nuclei (karyogamy). Karyogamy in a multinucleate LECA would promote the accumulation of duplications in all gene classes and promote genome expansion to its energetically permissible limits (Lane and Martin 2010) because unequal crossing between imprecisely paired homologous chromosomes following karyogamy generates duplications (Ohno 1970; Scannell et al. 2006; Hittinger and Carroll 2007; Van De Peer 2009). At the origin of meiotic recombination, chromosome pairing and segregation cannot have been perfect from the start; the initial state was likely error-prone, generating nuclei with aberrant gene copies, aberrant chromosomes, and even aberrant chromosome numbers. In cells with a single nucleus, such variants would have been lethal; in multinucleate (syncytial or coenocytic) organisms, defective nuclei can complement each other through mRNA in the cytosol (Garg and Martin 2016). Multinucleate forms are present throughout eukaryotic lineages (fig. 5), and ancestral reconstruction of nuclear organization clearly indicates that LECA itself was multinucleate (fig. 5 and supplementary fig. 1, Supplementary Material online). The multinucleate state enables the accumulation of duplications in the incipient eukaryotic lineage in a mechanistically nonadaptive manner, whereby duplications are implicated in the evolution of complexity (Ohno 1970; Scannell et al. 2006; Hittinger and Carroll 2007; Van De Peer 2009), as observed in the animal lineage (fig. 1). The syncytial state presents a viable intermediate state in the transition from prokaryote to eukaryote genetics.

Conclusion

Serial transfers of mitochondrial DNA to the chromosomes of the host are not only a mechanism of gene duplication, they are a form of endosymbiont genome duplication in which an original copy is retained in the organelle and remains functional. Gene duplications in LECA support an early origin of mitochondria and record the onset of the eukaryotic gene duplication process, a hallmark of genome evolution in mitosing cells (Ohno 1970; Scannell et al. 2006; Hittinger and Carroll 2007; Van De Peer 2009; Treangen and Rocha 2011).

Materials and Methods

Protein Clustering and Tree Reconstruction for Gene Duplication Inferences

Protein sequences for 150 eukaryotic genomes were downloaded from NCBI, Ensembl Protists, and JGI (see supplementary data 1 for detailed species composition). To construct gene families, we performed an all-vs-all BLAST (Altschul et al. 1997) of the eukaryotic proteins and selected the reciprocal best BLAST hits with e-value $\leq 10^{-10}$. The protein pairs were aligned with the Needleman–Wunsch algorithm (Rice et al. 2000) and the pairs with global identity values <25% were discarded. The retained global identity pairs were used to construct gene families with the Markov clustering algorithm (Enright et al. 2002) (version 12-068) with default parameters. Because in this study we were interested in gene duplications, we considered only the gene families with multiple gene copies in at least two eukaryotic genomes. Our criteria retained a total of 24,571 multicopy gene families.

Protein-sequence alignments for the individual eukaryotic multicopy gene families were generated using MAFFT (Katoh 2002), with the iterative refinement method that incorporates local pairwise alignment information (L-INS-i, version 7.130). The alignments were used to reconstruct maximum likelihood trees with IQ-tree (Nguyen et al. 2015), using default settings (version 1.6.5), and the trees were rooted with MAD (Tria et al. 2017) (supplementary data 2).

Inference of Gene Duplication

Gene duplications were inferred from gene trees by assigning duplication events to internal nodes in the rooted topologies. Given a rooted gene tree with n leaves, let S be the set of species labels for the leaves. For the case of paralogous gene trees, there is at least one leaf pair, a and b, such that $s_a = s_b$. Assigning a gene duplication to the last common ancestor of the pair a and b corresponds to the evolutionary scenario that minimizes paralog losses in the gene tree. For each rooted gene tree, we performed pairwise comparisons of all leaf pairs with identical species labels to infer all the internal nodes corresponding to gene duplications using the minimal loss criterion for each leaf pair. Note that, this approach considers

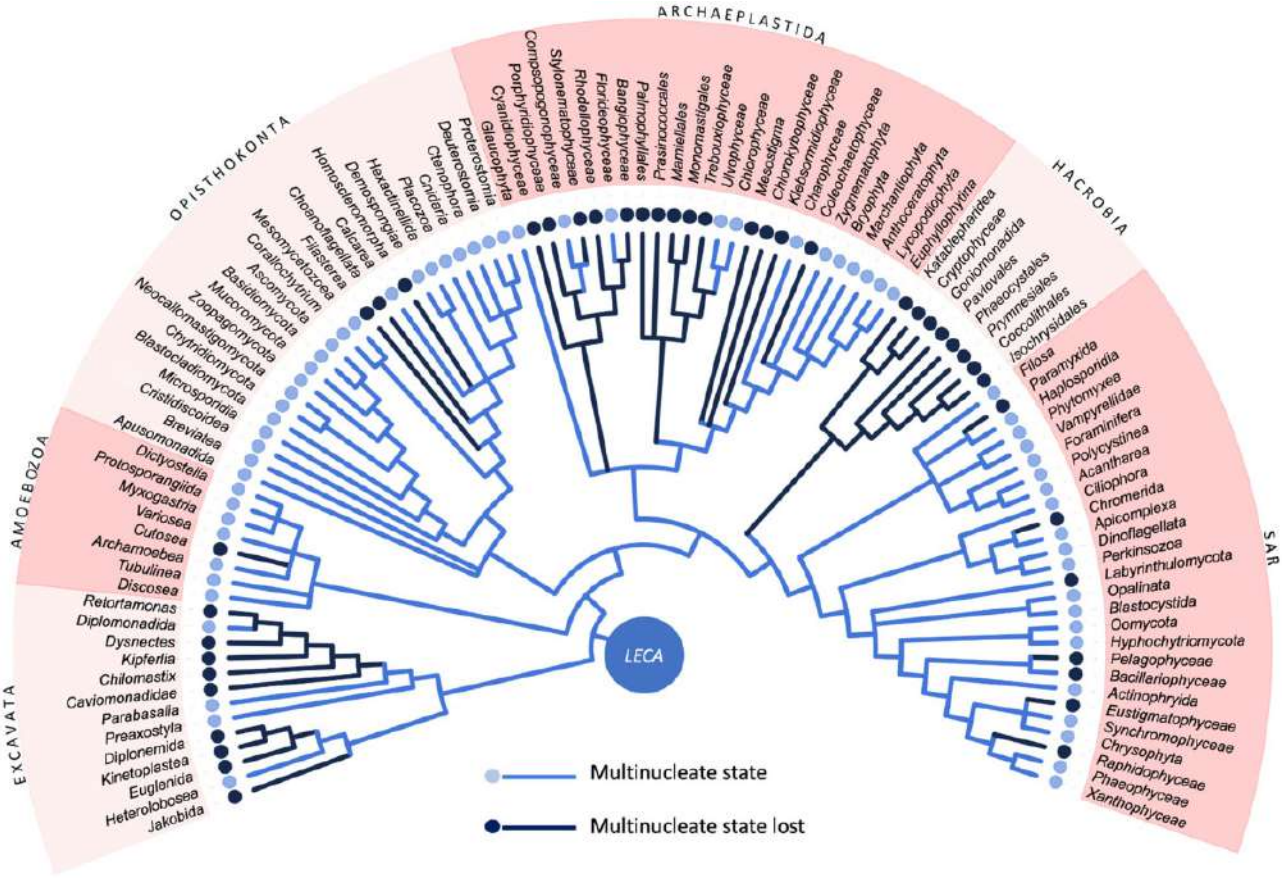


Fig. 5.—Ancestral state reconstruction for nuclear organization in eukaryotes. Presence and absence of the multinucleate state in members of the respective group are indicated. Resolution of the branches (polytomy vs. dichotomy) does not alter the outcome of the ancestral state reconstruction, nor does position of the root on the branches leading to Amoebozoa, Excavata, or Opisthokonta. LECA was a multinucleate, syncytial cell, not uninucleate (see supplementary fig. 1, Supplementary Material online). Together with mitochondrion and sex, the multinucleate state is ancestral to eukaryotes and fostered accumulation of duplications (see text).

the possibility of multiple gene duplications per gene tree (supplementary fig. 2, Supplementary Material online). We summarized the gene duplication inferences from all gene trees by evaluating the distribution of descendant paralogs across the eukaryotic supergroups for each gene duplication event (fig. 2).

The inferences of gene duplications in the present work are based on trees that were rooted with MAD (Tria et al. 2017). A recent comparison of MAD with other methods showed that MAD performs better than other rooting methods currently in use (Wade et al. 2020).

Inference for the Origin of Eukaryotic Duplicates

For identification of homologs in prokaryotes, we used all protein-coding genes from 5,656 prokaryotic genomes downloaded from RefSeq (Pruitt et al. 2007) (see supplementary data 3) and compared them against eukaryotic protein-coding genes using Diamond (Buchfink et al. 2015) to perform sequence searches with the "more-sensitive" parameter. A eukaryotic gene family was considered to have homologs in prokaryotes if at least one gene of the eukaryotic family had a significant hit against a prokaryotic gene (e-value $<10^{-10}$ and local identity $\geq$25%). Gene families with homologs only in archaeal genomes were considered as genes of archaeal origin and similarly for bacteria. Gene families with significant hits in both archaea and bacteria (universal) could have originated from either archaea or bacteria.

We purposefully avoided using trees to inferring the origin of eukaryotic genes because of low levels of sequence conservation entailing a large number of prokaryotic homologs. Note, however, that we reconstructed trees for the subset of eukaryote–prokaryote genes with sufficient sequence conservation (see below). We found that the presence–absence of homologs across prokaryotic taxa remarkably recapitulates the distribution of prokaryotic sisters derived from phylogenetic trees serving, thus, as a validation of our approach (supplementary table 5).

Prokaryote–Eukaryote Protein Clustering and Tree Reconstruction

To assemble a data set of conserved genes for phylogenies linking prokaryotes and eukaryotes, eukaryotic, archaeal, and bacterial protein sequences were first clustered separately before homologous clusters between eukaryotes and prokaryotes were identified as described (Ku et al. 2015). Eukaryotic sequences for the 150 genomes (supplementary data 1) were clustered with MCL (Enright et al. 2002) using global identities from best reciprocal BLAST (Altschul et al. 1997) hits for protein pairs with e-value $\leq 10^{-10}$ and global identity $\geq 40\%$. The clusters with genes distributed in more than one eukaryotic genome were retained. Similarly, prokaryotic protein sequences from 5,655 genomes (see supplementary data 3, except for MK-D1 for which the genome was unavailable by the time the data were compiled) were clustered using the best reciprocal BLAST for protein pairs with e-value $\leq 10^{-10}$ and global identity $\geq 25\%$, for archaea and bacteria separately. The resulting clusters with gene copies in at least five prokaryotic genomes were retained. The most universally distributed clusters comprise 20–40 proteins, the majority of which are involved in translation (supplementary fig. 4, Supplementary Material online). Eukaryotic and prokaryotic clusters were merged using the reciprocal best cluster procedure. We merged a eukaryotic cluster with a prokaryotic cluster if $\geq 50\%$ of the eukaryotic sequences in the cluster have their best reciprocal BLAST hit in the same prokaryotic cluster and vice versa (cut-offs: e-value $\leq 10^{-10}$ and local identity $\geq 30\%$). We refer to the merged cluster as eukaryotic–prokaryotic cluster (EPC).

Protein-sequence alignments for 2,575 EPCs were generated using MAFFT (Katoh 2002) (L-INS-i, version 7.130). The alignments were used to reconstructed maximum-likelihood trees with IQ-tree (Nguyen et al. 2015) (version 1.6.5) employing default settings (supplementary data 4).

Tests for Eukaryote Monophyly

For 475 gene trees where eukaryotes were not recovered as monophyletic, we conducted the Shimodaira–Hasegawa (Shimodaira and Hasegawa 1999) (SH), Kishino–Hasegawa (Kishino and Hasegawa 1989) (KH), and approximately unbiased (AU) test (Shimodaira 2002) to determine whether the observed nonmonophyly was statistically significant. We reconstructed trees constraining eukaryotic sequences to be monophyletic, but not imposing any other topological constraint, using FastTree (Price et al. 2010) (version 2.1.10 SSE3) and recording all trees explored during the tree search with the "-log" parameter (supplementary data 5). The sample of monophyletic trees was used as input in IQ-tree (Nguyen et al. 2015) (version 2.0.3; parameter: "-zb 100000 –au") to perform the SH, KH, and AU tests against the unconstrained tree (nonmonophyletic). If the best-constrained tree did not show significant difference relative to the unconstrained tree ($P < 0.05$), then we considered that eukaryotic monophyly cannot be rejected.

Inference of Prokaryotic Sisters

To infer prokaryotes sisters to eukaryotes in the gene trees we used the unconstrained tree if eukaryotes were recovered as monophyletic and the constrained tree if eukaryotes were not recovered as monophyletic, since the SH test did not reject eukaryote monophyly for any gene tree (see main text). Note that in unrooted trees for which eukaryotes are monophyletic, the prokaryotic side of the tree is bisected by one internal node into two prokaryotic subclades, each subclade being the potential sister to eukaryotes (see fig. 4*a*). We considered the prokaryotic subclade with the smallest number of leaves for our inferences of sister-relations and the prokaryotic phyla present in the sister clade and outgroup clade was recorded for each tree. The sister clades were scored as a "pure" sister when only a single prokaryotic phylum was present in the clade or as "mixed" sister when more than one phylum was present.

Ancestral Reconstruction of Eukaryotic Nuclear Organization

Ancestral state reconstructions were performed on the basis of a morphological character matrix, using maximum parsimony as implemented in Mesquite 3.6 (https://www.mesquiteproject.org/, accessed June 2019). The reference eukaryotic phylogeny includes 106 taxa (ranging from genus to phylum level) to reflect the relations within the eukaryotes and reduce taxonomic redundancy. The phylogeny includes members of six supergroups: Amoebozoa (Mycetozoa), Archaeplastida, Excavata, Hacrobia, Opisthokonta, and SAR, and was constructed by combining branches from previous studies (Burki et al. 2010; Yoon et al. 2010; Adl et al. 2012; Powell and Letcher 2014; Burki et al. 2016; Cavalier-Smith et al. 2016; Derelle et al. 2016; Spatafora et al. 2016; Yang et al. 2016; Archibald et al. 2017; Krabberød et al. 2017; McCarthy and Fitzpatrick 2017; Roger et al. 2017; Spatafora et al. 2017; Bass et al. 2018; Cavalier-Smith et al. 2018; Tedersoo et al. 2018; Irwin et al. 2019). The nuclear organization for each taxon was coded as 0 for nonmultinucleate, 1 for multinucleate or 0/1 if ambiguous according to the literature (Byers 1979; Willumsen et al. 1987; Barthel and Detmer 1990; Daniels and Pappas 1994; Walker et al. 2006; Steiner 2010; Yoon et al. 2010; Adl et al. 2012; Niklas et al. 2013; Maciver 2016; Spatafora et al. 2016; Archibald et al. 2017; Bloomfield et al. 2019) (supplementary data 6). In order to account for uncertainties of lineage relations among eukaryotes, we used a set of phylogenies with alternative root positions (Vossbrinck et al. 1987; Stechmann and Cavalier-Smith 2002; Katz and Grant 2015) (altogether a total of 15 different roots) as well as the consideration of polytomies for debated branches (supplementary data 6). All ancestral state reconstruction

rendered LECA as multinucleated, with no ambiguity. Ambiguous reconstructions, however, were observed within supergroups in some topologies but did not pose ambiguity to the reconstructed state in LECA.

Supplementary Material

Supplementary data are available at *Genome Biology and Evolution* online.

Acknowledgments

We thank the European Research Council (Grant No. 666053), the Volkswagen Foundation (Grant No. 93 046), and the Moore Simons Initiative on the Origin of the Eukaryotic Cell (Grant No. 9743) for financial support. We also thank Damjan Franjević (Department of Biology, University of Zagreb, Croatia) for helpful discussions.

Author Contributions

All authors conceived and designed the study. J.B. and J.S. prepared the data sets with contribution from all the authors. F.D.K.T. performed gene duplication inferences, functional annotation of genes, and the tests for eukaryotic monophyly. J.B. and F. N. performed the analyses of eukaryotic sisters. J.S. compiled the eukaryotic phylogenies and performed ancestral state reconstructions. All authors wrote the paper.

Data Availability

Supplementary tables and data used in this study are available under the link https://doi.org/10.6084/m9.figshare.12249260.

Code Availability

Custom Matlab scripts used to perform data analysis are available upon request.

Literature Cited

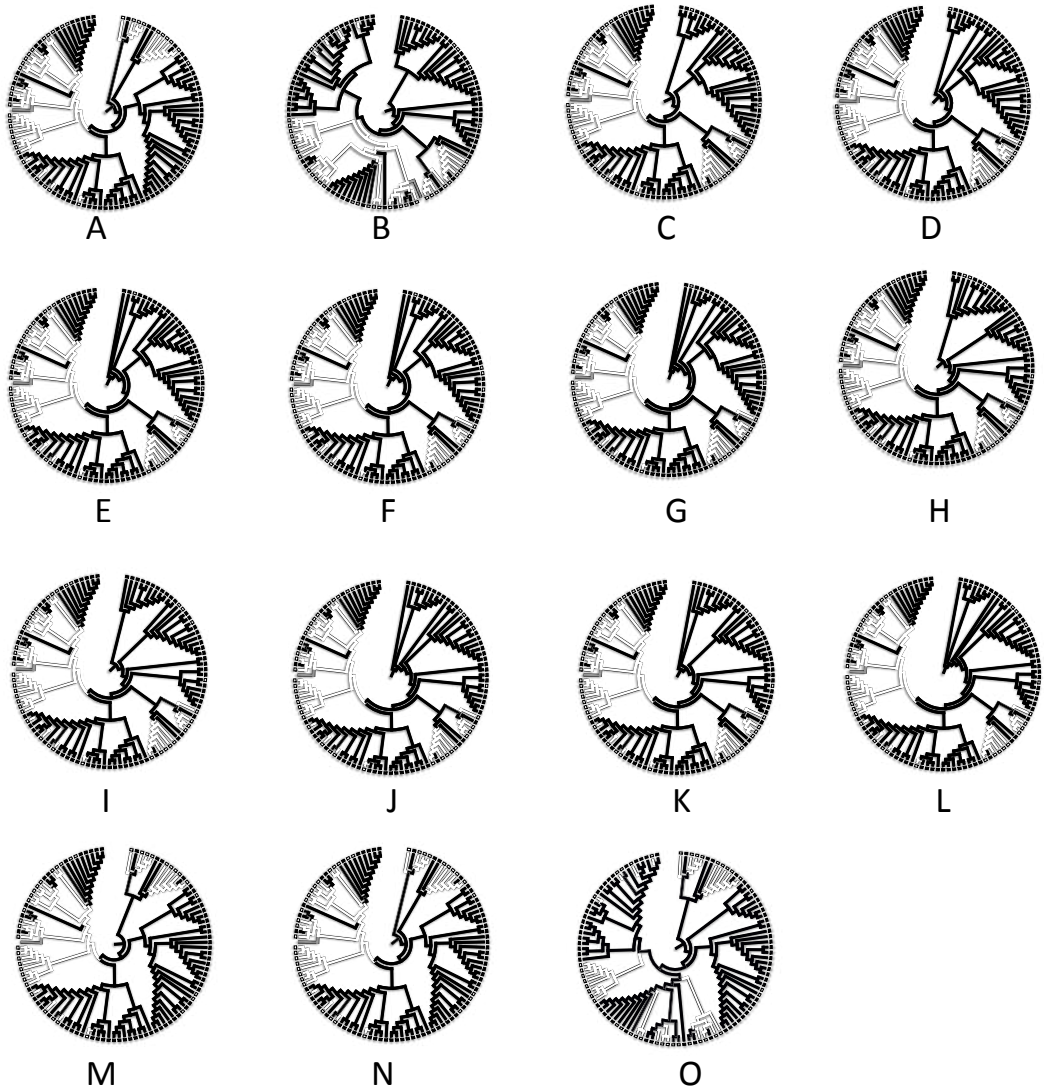
Adl SM, et al. 2012. The revised classification of eukaryotes. J Eukaryot Microbiol. 59(5):429–493.
Albalat R, Cañestro C. 2016. Evolution by gene loss. Nat Rev Genet. 17(7):379–391.
Allen JF. 2015. Why chloroplasts and mitochondria retain their own genomes and genetic systems: colocation for redox regulation of gene expression. Proc Natl Acad Sci U S A. 112(33):10231–10238.
Altschul SF, et al. 1997. Blast and Psi-Blast: protein database search programs. Nucleic Acid Res. 25:2289–4402.
Andersson JO, et al. 2003. Phylogenetic analyses of diplomonad genes reveal frequent lateral gene transfers affecting eukaryotes. Curr Biol. 13:94–104.
Archibald JM, et al. 2017. Handbook of the protists. Cham: Springer Nature.
Barlow LD, Nývltová E, Aguilar M, Tachezy J, Dacks JB. 2018. A sophisticated, differentiated Golgi in the ancestor of eukaryotes. BMC Biol. 16(1):27.
Barthel D, Detmer A. 1990. The spermatogenesis of *Halichondria panicea* (Porifera, Demospongiae). Zoomorphology 110:9–15.
Bass D, et al. 2018. Clarifying the relationships between microsporidia and cryptomycota. J Eukaryot Microbiol. 65(6):773–782.
Betts HC, et al. 2018. Integrated genomic and fossil evidence illuminates life's early evolution and eukaryote origin. Nat Ecol Evol. 2:1556–1562.
Bloomfield G, et al. 2019. Triparental inheritance in *Dictyostelium*. Proc Natl Acad Sci U S A. 116(6):2187–2192.
Booth A, Doolittle WF. 2015. Eukaryogenesis, how special really? Proc Natl Acad Sci U S A. 112(33):10278–10285.
Buchfink B, Xie C, Huson DH. 2015. Fast and sensitive protein alignment using DIAMOND. Nat Methods. 12(1):59–60.
Burki F, et al. 2010. Evolution of Rhizaria: new insights from phylogenomic analysis of uncultivated protists. BMC Evol Biol. 10:377.
Burki F, et al. 2016. Untangling the early diversification of eukaryotes: a phylogenomic study of the evolutionary origins of Centrohelida, Haptophyta and Cryptista. Proc R Soc Lond B. 283:20152802.
Byers TJ. 1979. Growth, reproduction, and differentiation in Acanthamoeba. Int Rev Cytol. 61:283–338.
Cavalier-Smith T, Chao EE. 2020. Multidomain ribosomal protein trees and the planctobacterial origin of neomura (eukaryotes, archaebacteria). Protoplasma 257(3):621–753.
Cavalier-Smith T, et al. 2016. 187-gene phylogeny of protozoan phylum Amoebozoa reveals a new class (Cutosea) of deep-branching, ultrastructurally unique, enveloped marine Lobosa and clarifies amoeba evolution. Mol Phylogenet Evol. 99:275–296.
Cavalier-Smith T, et al. 2018. Multigene phylogeny and cell evolution of chromist infrakingdom Rhizaria: contrasting cell organisation of sister phyla Cercozoa and Retaria. Protoplasma 255(5):1517–1574.
Cavalier-Smith T. 1975. The origin of nuclei and of eukaryotic cells. Nature 256:463–468.
Cavalier-Smith T. 2002. The phagotrophic origin of eukaryotes and phylogenetic classification of Protozoa. Int J Syst Evol Microbiol. 52(Pt 2):297–354.
Daniels EW, Pappas GD. 1994. Reproduction of nuclei in *Pelomyxa palustris*. Cell Biol Int. 18(8):805–812.
de Duve C. 2007. The origin of eukaryotes: a reappraisal. Nat Rev Genet. 8(5):395–403.
Derelle R, et al. 2016. Phylogenomic framework to study the diversity and evolution of Stramenopiles (= Heterokonts). Mol Biol Evol. 33(11):2890–2898.
Dolezal P, Likic V, Tachezy J, Lithgow T. 2006. Evolution of the molecular machines for protein import into mitochondria. Science 313(5785):314–318.
Doolittle FW. 1998. You are what you eat: a gene transfer ratchet could account for bacterial genes in eukaryotic nuclear genomes. Trends Genet. 14(8):307–311.
Embley T, Martin W. 2006. Eukaryotic evolution, changes and challenges. Nature 440(7084):623–630.
Eme L, et al. 2017. Archaea and the origin of eukaryotes. Nat Rev Microbiol. 15(12):711–723.
Enright AJ, Van Dongen S, Ouzounis CA. 2002. An efficient algorithm for large-scale detection of protein families. Nucleic Acids Res. 30(7):1575–1584.
Esser C, et al. 2004. A genome phylogeny for mitochondria among alpha-proteobacteria and a predominantly eubacterial ancestry of yeast nuclear genes. Mol Biol Evol. 21(9):1643–1660.
Esser C, Martin W, Dagan T. 2007. The origin of mitochondria in light of a fluid prokaryotic chromosome model. Biol Lett. 22:180–184.
French SL, Santangelo TJ, Beyer AL, Reeve JN. 2007. Transcription and translation are coupled in Archaea. Mol Biol Evol. 24(4):893–895.
Gabaldón T. 2018. Relative timing of mitochondrial endosymbiosis and the "pre-mitochondrial symbioses" hypothesis. IUBMB Life. 70(12):1188–1196.

Garg SG, Martin WF. 2016. Mitochondria, the cell cycle, and the origin of sex via a syncytial eukaryote common ancestor. Genome Biol. Evol. 8:1950–1970.
Gould SB, Garg SG, Martin WF. 2016. Bacterial vesicle secretion and the evolutionary origin of the eukaryotic endomembrane system. Trends Microbiol. 24(7):525–534.
Gray MW. 2014. The pre-endosymbiont hypothesis: a new perspective on the origin and evolution of mitochondria. Cold Spring Harb Perspect Biol. 6:a016097.
Hampl V, Čepička I, Eliáš M. 2019. Was the mitochondrion necessary to start eukaryogenesis? Trends Microbiol. 27(2):96–104.
Hampl V, et al. 2009. Phylogenomic analyses support the monophyly of Excavata and resolve relationships among eukaryotic 'supergroups'. Proc Natl Acad Sci U S A. 106(10):3859–3864.
He D, et al. 2014. An alternative root for the eukaryote tree of life. Curr Biol. 24(4):465–470.
Hittinger CT, Carroll SB. 2007. Gene duplication and the adaptive evolution of a classic genetic switch. Nature 449(7163):677–681.
Imachi H, et al. 2020. Isolation of an archaeon at the prokaryote-eukaryote interface. Nature 577(7791):519–525.
Irwin NA, et al. 2019. Phylogenomics supports the monophyly of the Cercozoa. Mol Phylogenet Evol. 130:416–423.
Javaux EJ, Lepot K. 2018. The Paleoproterozoic fossil record: implications for the evolution of the biosphere during Earth's middle-age. Earth Sci Rev. 176:68–86.
Jeffroy O, Brinkmann H, Delsuc F, Philippe H. 2006. Phylogenomics: the beginning of incongruence? Trends Genet. 22(4):225–231.
Katoh K, Misawa K, Kuma K-I, Miyata T. 2002. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acids Res. 30(14):3059–3066.
Katz LA, Grant JR. 2015. Taxon-rich phylogenomic analyses resolve the eukaryotic tree of life and reveal the power of subsampling by sites. Syst Biol. 64(3):406–415.
Keeling PJ, Palmer LD. 2008. Horizontal gene transfer in eukaryotic evolution. Nat Rev Genet. 9(8):605–618.
Kishino H, Hasegawa M. 1989. Evaluation of the maximum-likelihood estimate of the evolutionary tree topologies from DNA sequence data, and the branching order in Hominoidea. J Mol Evol. 29(2):170–179.
Kollmar M. 2016. Fine-tuning motile cilia and flagella: evolution of the dynein motor proteins from plants to humans at high resolution. Mol Biol Evol. 33(12):3249–3267.
Krabberød AK, et al. 2017. Single cell transcriptomics, mega-phylogeny, and the genetic basis of morphological innovations in Rhizaria. Mol Biol Evol. 34(7):1557–1573.
Ku C, et al. 2015. Endosymbiotic origin and differential loss of eukaryotic genes. Nature 524(7566):427–432.
Ku C, Martin WF. 2016. A natural barrier to lateral gene transfer from prokaryotes to eukaryotes revealed from genomes: the 70% rule. BMC Biol. 14(1):89.
Lake JA, Rivera MC. 1994. Was the nucleus the first endosymbiont? Proc Natl Acad Sci U S A. 91(8):2880–2881.
Lane N, Martin W. 2010. The energetics of genome complexity. Nature 467(7318):929–934.
Leger MM, et al. 2018. Demystifying eukaryote lateral gene transfer. Bioessays 40(5):e1700242.
López-García P, Moreira D. 1999. Metabolic symbiosis at the origin of eukaryotes. Trends Biochem Sci. 24:88–93.
López-García P, Moreira G. 2015. Open questions on the origin of eukaryotes. Trends Ecol Evol. 30(11):697–708.
Maciver SK. 2016. Asexual amoebae escape Muller's ratchet through polyploidy. Trends Parasitol. 32(11):855–862.
Makarova KS, et al. 2005. Ancestral paralogs and pseudoparalogs and their role in the emergence of the eukaryotic cell. Nucleic Acids Res. 33(14):4626–4638.
Margulis L, Chapman M, Guerrero R, Hall J. 2006. The last eukaryotic common ancestor (LECA): acquisition of cytoskeletal motility from aerotolerant spirochetes in the Proterozoic Eon. Proc Natl Acad Sci U S A. 103(35):13080–13085.
Margulis L, et al. 2000. The chimeric eukaryote: origin of the nucleus from the karyomastigont in amitochondriate protists. Proc Natl Acad Sci U S A. 97(13):6954–6959.
Martin W. 1999. Mosaic bacterial chromosomes: a chalenge en route to a tree of genomes. Bioessays 21:99–104.
Martin W. 2010. Evolutionary origins of metabolic compartmentalization in eukaryotes. Philos Trans R Soc Lond B Biol Sci. 365(1541):847–855.
Martin W, Brinkmann H, Savonna C, Cerff R. 1993. Evidence for a chimeric nature of nuclear genomes: eubacterial origin of eukaryotic glyceraldehyde-3-phosphate dehydrogenase genes. Proc Natl Acad Sci U S A. 90(18):8692–8696.
Martin W, et al. 2001. An overview of endosymbiotic models for the origins of eukaryotes, their ATP-producing organelles (mitochondria and hydrogenosomes), and their heterotrophic lifestyle. Biol Chem. 382(11):1521–1539.
Martin W, Müller M. 1998. The hydrogen hypothesis for the first eukaryote. Nature 392(6671):37–41.
Martin WF, Garg S, Zimorski V. 2015. Endosymbiotic theory for eukaryote origin. Philos Trans R Soc Lond B. 370:20140330.
Martin WF, Tielens AGM, Mentel M, Garg SG, Gould SB. 2017. The physiology of phagocytosis in the context of mitochondrial origin. *Microbiol*. Mol Biol Rev. 81:e00008–e00017.
McCarthy CG, Fitzpatrick DA. 2017. Multiple approaches to phylogenomic reconstruction of the fungal kingdom. Adv Genet. 100:211–266.
Mereschkowsky C. 1905. Über Natur und Ursprung der Chromatophoren im Pflanzenreiche. Biol Centralbl. 25:593–604. English translation in Martin W, Kowallik KV. 1999. Annotated English translation of Mereschkowsky's 1905 paper 'Über Natur und Ursprung der Chromatophoren im Pflanzenreiche '. Eur J Phycol., 34:287–295.
Nagies FSP, Brueckner J, Tria FDK, Martin WF. 2020. A spectrum of verticality across genes. PLoS Genet. 16(11):e1009200.
Nei M, Gu X, Sitnikova T. 1997. Evolution by birth and death process in multigene families of the vertebrate immune system. Proc Natl Acad Sci U S A. 94(15):7799–7806.
Nelson-Sathi S, et al. 2012. Acquisition of 1,000 eubacterial genes physiologically transformed a methanogen at the origin of Haloarchaea. Proc Natl Acad Sci U S A. 109(50):20537–20542.
Nguyen LT, Schmidt HA, Von Haeseler A, Minh BQ. 2015. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol Biol Evol. 32(1):268–274.
Niklas KJ, et al. 2013. The evo-devo of multinucleate cells, tissues, and organisms, and an alternative route to multicellularity. Evol Dev. 15(6):466–474.
Ohno S. 1970. Evolution by gene duplication. Heidelberg (Berlin): Springer.
Parfrey LW, et al. 2011. Estimating the timing of early eukaryotic diversification with multigene molecular clocks. Proc Natl Acad Sci U S A. 108:1364–13629.
Pittis AA, Gabaldón T. 2016. Late acquisition of mitochondria by a host with chimaeric prokaryotic ancestry. Nature 531(7592):101–104.
Poole AM, Gribaldo S. 2014. Eukaryotic origin: how and when was the mitochondrion acquired? Cold Spring Harb Perspect Biol. 6(12):a015990.
Portugez S, Martin WF, Hazkani-Covo E. 2018. Mosaic mitochondrial-plastid insertions into the nuclear genome show evidence of both non-homologous end joining and homologous recombination. BMC Evol Biol. 18(1):162.

Powell MJ, Letcher PM. 2014. 6 Chytridiomycota, Monoblepharidomycota, and Neocallimastigomycota. In: McLaughlin DJ, Spatafora JW, editors. 2nd ed. The Mycota Part VII A. Systematics and evolution. Heidelberg (Berlin): Springer. p. 141–175.
Price MN, Dehal PS, Arkin AP. 2010. FastTree 2 – approximately maximum-likelihood trees for large alignments. PLoS One 5(3):e9490.
Pruitt KD, Tatusova T, Maglott DR. 2007. NCBI reference sequences (RefSeq): a curated non-redundant sequence database of genomes, transcripts and proteins. Nucleic Acids Res. 35(Database issue):D61–D65.
Ren R, et al. 2016. Phylogenetic resolution of deep eukaryotic and fungal relationships using highly conserved low-copy nuclear genes. Genome Biol Evol. 8(9):2683–2701.
Rice P, et al. 2000. EMBOSS: the European Molecular Biology Open software suite. Trends Genet. 16(6):276–277.
Río Bártulos C, et al. 2018. Mitochondrial glycolysis in a major lineage of eukaryotes. Genome Biol Evol. 10(9):2310–2325.
Roger AJ, Muñoz-Gómez SA, Kamikawa R. 2017. The origin and diversification of mitochondria. Curr Biol. 27(21):R1177–R1192.
Scannell DR, Byrne KP, Gordon JL, Wong S, Wolfe KH. 2006. Multiple rounds of speciation associated with reciprocal gene loss in polyploid yeasts. Nature 440(7082):341–345.
Shimodaira H. 2002. An approximately unbiased test of phylogenetic tree selection. Syst Biol. 51(3):492–508.
Shimodaira H, Hasegawa M. 1999. Multiple comparisons of log-likelihoods with applications to phylogenetic inference. Mol Biol Evol. 16:1114–1116.
Spang A, et al. 2015. Complex archaea that bridge the gap between prokaryotes and eukaryotes. Nature 521(7551):173–179.
Spatafora JW, et al. 2016. A phylum-level phylogenetic classification of zygomycete fungi based on genome-scale data. Mycologia 108(5):1028–1046.
Spatafora JW, et al. 2017. The fungal tree of life: from molecular systematics to genome-scale phylogenies. Microbiol Spectr. 5(5):1–32.
Speijer D, Lukeš J, Eliáš M. 2015. Sex is a ubiquitous, ancient, and inherent attribute of eukaryotic life. Proc Natl Acad Sci U S A. 112(29):8827–8834.
Stechmann A, Cavalier-Smith T. 2002. Rooting the eukaryote tree by using a derived gene fusion. Science 297(5578):89–91.
Steiner JM. 2010. Technical notes: growth of *Cyanophora paradoxa*. J Endoc Cell Res. 20:62–67.
Tedersoo L, et al. 2018. High-level classification of the fungi and a tool for evolutionary ecological analyses. Fungal Div. 90:135–159.
Timmis JN, Ayliff MA, Huang CY, Martin W. 2004. Endosymbiotic gene transfer: organelle genomes forge eukaryotic chromosomes. Nat Rev Genet. 5(2):123–135.
Treangen TJ, Rocha EPC. 2011. Horizontal transfer, not duplication, drives the expansion of protein families in prokaryotes. PLoS Genet. 7(1):e1001284.
Tria FDK, Landan G, Dagan T. 2017. Phylogenetic rooting using minimal ancestor deviation. Nat Ecol Evol. 1:0193.
Tromer EC, van Hooff JJE, Kops GJPL, Snel B. 2019. Mosaic origin of the eukaryotic kinetochore. Proc Natl Acad Sci U S A. 116(26):12873–12882.
Van De Peer Y, Maere S, Meyer A. 2009. The evolutionary significance of ancient genome duplications. Nat Rev Genet. 10(10):725–732.
Vossbrinck CR, et al. 1987. Ribosomal RNA sequence suggests microsporidia are extremely ancient eukaryotes. Nature 326(6111):411–414.
Vosseberg J, et al. 2021. Timing the origin of eukaryotic cellular complexity with ancient duplications. Nat Ecol Evol. 5(1):92–100.
Wade T, et al. 2020. Assessing the accuracy of phylogenetic rooting methods on prokaryotic gene families. PLoS One 15(5):e0232950–e0233022.
Walker G, et al. 2006. Ultrastructural descripton of *Breviata anathema*, n. gen., n. sp., the organism previously studied as "*Mastigamoeba invertens*". J Eukaryot Microbiol. 53(2):65–78.
Wallin IE. 1925. On the nature of mitochondria. IX. Demonstration of the bacterial nature of mitochondria. Am J Anat. 36:131–139.
Willumsen NB, et al. 1987. A multinucleate amoeba, *Parachaos zoochlorellae* (Willumsen 1982) comb, nov., and a proposed division of the genus *Chaos* into the Genera *Chaos* and *Parachaos* (Gymnamoebia, Amoebidae). Archiv Protist. 134:303–313.
Yang EC, et al. 2016. Divergence time estimates and the evolution of major lineages in the florideophyte red algae. Sci Rep. 6:21361.
Yoon HS, et al. 2010. Evolutionary history and taxonomy of red algae. In: Seckbach, JChapman, DJ, editors. Red algae in genomic age. Dordrecht: Springer. p. 27–45.
Zachar I, Szathmáry E. 2017. Breath-giving cooperation: critical review of origin of mitochondria hypotheses. Biol Direct. 12:19.
Zaremba-Niedzwiedzka K, et al. 2017. Asgard archaea illuminate the origin of eukaryotic cellular complexity. Nature 541(7637):353–358.
Zmasek CM, Godzik A. 2011. Strong functional patterns in the evolution of eukaryotic genomes revealed by the reconstruction of ancestral protein domain repertoires. Genome Biol. 12(1):R4.

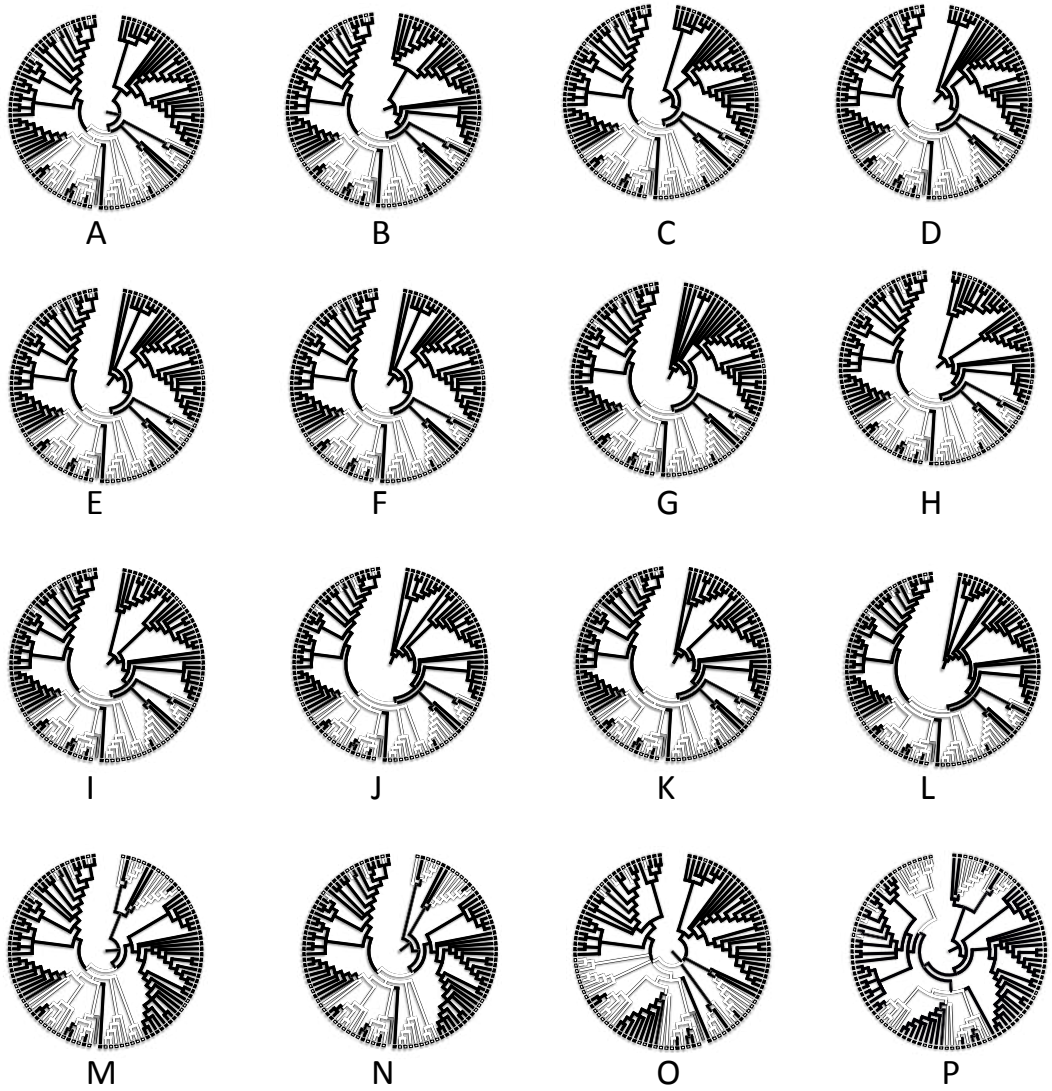
Associate editor: Ellen Pritham

Supplemental Figures



A) Dicho_01 only dichotomy allowed, rootAmorphaea (Opisthokonta+Amoebozoa)
B) Dicho_02 only dichotomy allowed, root Opisthokonta
C) Dicho_03 only dichotomy allowed, root Amoebozoa
D) Dicho_04 only dichotomy allowed, root Evosea (within Amoebozoa)
E) Dicho_05 only dichotomy allowed, root Tubulinea (within Amoebozoa)
F) Dicho_06 only dichotomy allowed, root Discosea (within Amoebozoa)
G) Dicho_07 only dichotomy allowed, root Archamoebea (within Amoebozoa: Evosea)
H) Dicho_08 only dichotomy allowed, root Holozoa
I) Dicho_09 only dichotomy allowed, root Holomycota (Fungi s.l.)
J) Dicho_10 only dichotomy allowed, root Microsporidia (within Fungi)
K) Dicho_11 only dichotomy allowed, root Cristidiscoidea (within Fungi)
L) Dicho_12 only dichotomy allowed, root Blastocladiomycota (within Fungi)
M) Dicho_13 only dichotomy allowed, root Excavata (monophyletic)
N) Dicho_14 only dichotomy allowed, root Discoba (within Excavata)
O) P) Dicho_duplo only dichotomy allowed, rooted with Excavata, this paper's topology

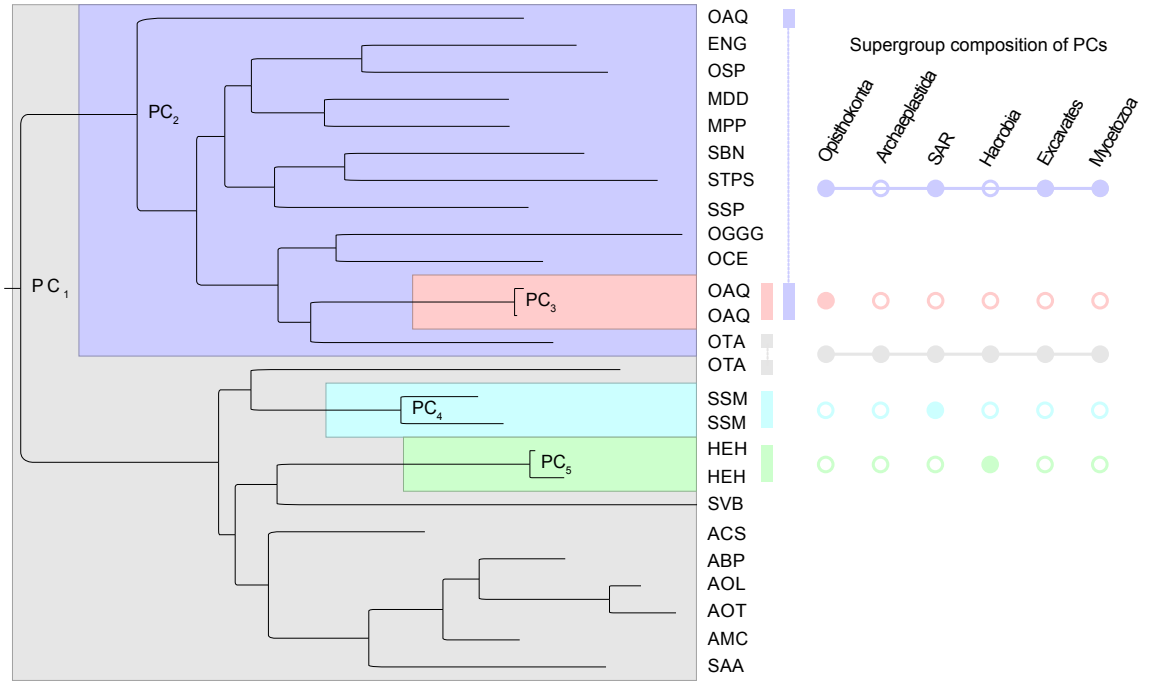
MULTINUCLEATE REPRESENTATIVES KNOWN
NO MULTINUCLEATE REPRESENTATIVES



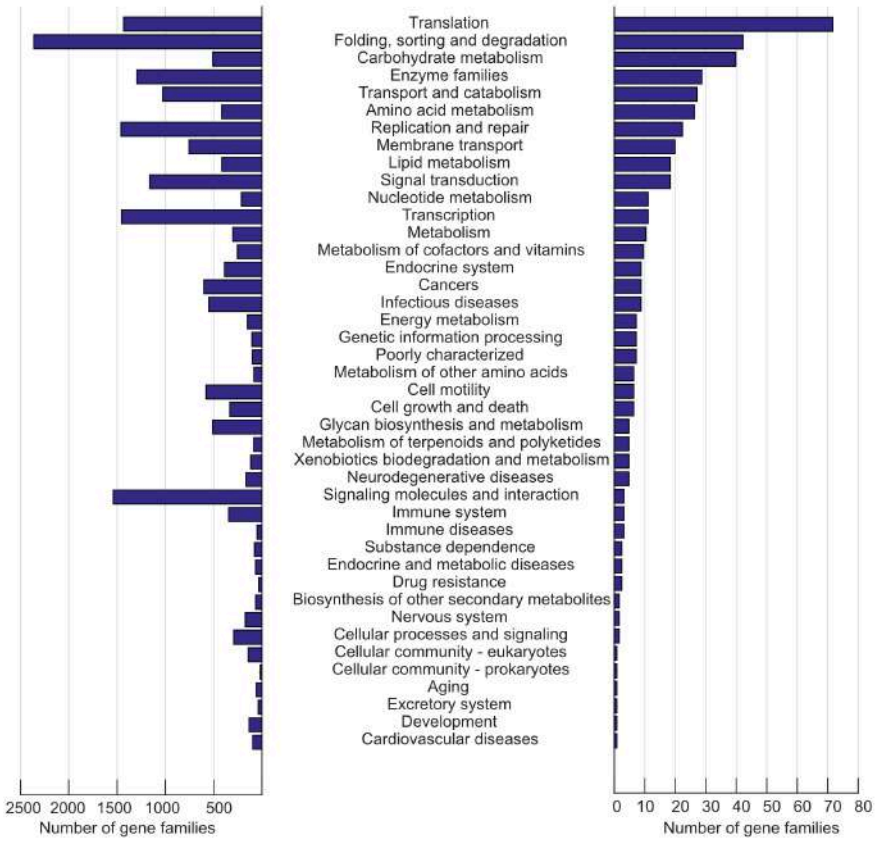
A) Poly_01 polytomy allowed, rootAmorphaea (Opisthokonta+Amoebozoa)
B) Poly_02 polytomy allowed, root Opisthokonta
C) Poly_03 polytomy allowed, root Amoebozoa
D) Poly_04 polytomy allowed, root Evosea (within Amoebozoa)
E) Poly_05 polytomy allowed, root Tubulinea (within Amoebozoa)
F) Poly_06 polytomy allowed, root Discosea (within Amoebozoa)
G) Poly_07 polytomy allowed, root Archamoebea (within Amoebozoa: Evosea)
H) Poly_08 polytomy allowed, root Holozoa
I) Poly_09 polytomy allowed, root Holomycota (Fungi s.l.)
J) Poly_10 polytomy allowed, root Microsporidia (within Fungi)
K) Poly_11 polytomy allowed, root Cristidiscoidea (within Fungi)
L) Poly_12 polytomy allowed, root Blastocladiomycota (within Fungi)
M) Poly_13 polytomy allowed, root Excavata (monophyletic)
N) Poly_14 polytomy allowed, root Discoba (within Excavata)
O) Poly_00_max polytomy allowed, not rooted
P) Poly_duplo polytomy allowed, rooted with Excavata, this paper's topology

MULTINUCLEATE REPRESENTATIVES KNOWN
NO MULTINUCLEATE REPRESENTATIVES

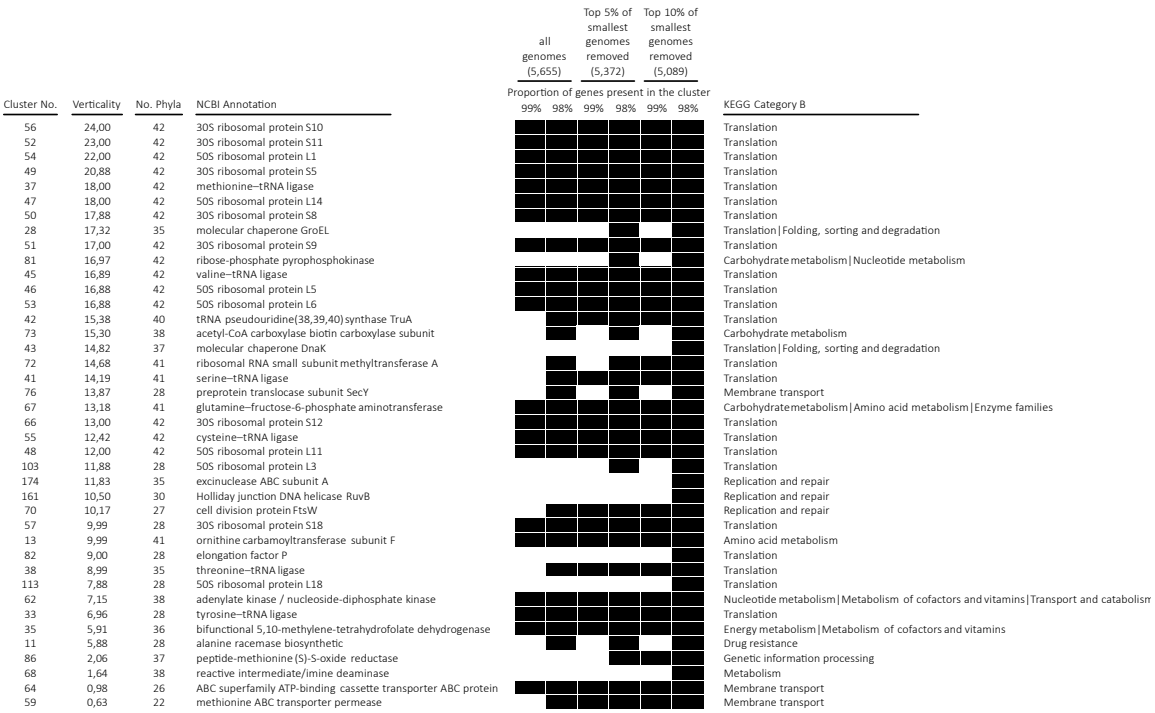
Supplemental Figure 1: Ancestral reconstruction of nuclear organization in eukaryotes. The ancestral reconstruction of a LECA as multinucleated organism is consistent for all 31 phylogenies of 106 eukaryotic taxa tested in this study (see **Supplemental Data 6**).



Supplemental Figure 2: Gene duplication inference for one example gene tree. The figure shows a maximum-likelihood tree reconstructed from aldolase protein sequences. The pairwise comparisons of paralogs with identical species labels (colored bars on the right) enables the inference of internal duplication nodes in the gene tree. Genes descending from duplication nodes form paralogous clades (PC, colored clades), with varying species composition (filled circles on the right). PC3, PC4 and PC5 are exclusive to species affiliated to single supergroups, whereas PC1 and PC2 are PCs with species from multiple supergroups. The first character in the leaf labels indicate the supergroup affiliation of the species. O for Opisthokonta, H for Hacrobia, S for SAR, M for Mycetozoa and A for Archaeplastida. The remaining characters in the leaf labels refer to species names (see full species names in **Supplemental Data 1**).



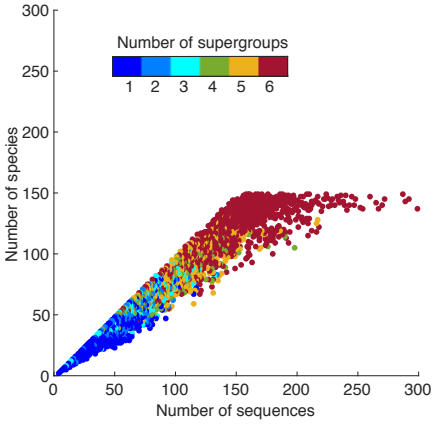
Supplemental Figure 3: Distribution functional categories across eukaryotic multi-copy gene families. The eukaryotic protein sequences were compared against the KEGG database with Diamond to align protein sequences. We retained the results with e-value < 10^{-10} and >=25% of identical residues relative to the eukaryotic query. Functional annotations were assigned to the eukaryotic gene families according to the best subject sequence in KEGG as judged by minimum e-value criterion from Diamond searches. Left panel: Eukaryotic gene families without duplications in LECA but with duplications in more recent ancestors. Right panel: Eukaryotic gene families with duplications in LECA.



Cluster No.	Verticality	No. Phyla	NCBI Annotation	KEGG Category B
56	24,00	42	30S ribosomal protein S10	Translation
52	23,00	42	30S ribosomal protein S11	Translation
54	22,00	42	50S ribosomal protein L1	Translation
49	20,88	42	30S ribosomal protein S5	Translation
37	18,00	42	methionine--tRNA ligase	Translation
47	18,00	42	50S ribosomal protein L14	Translation
50	17,88	42	30S ribosomal protein S8	Translation
28	17,32	35	molecular chaperone GroEL	Translation|Folding, sorting and degradation
51	17,00	42	30S ribosomal protein S9	Translation
81	16,97	42	ribose-phosphate pyrophosphokinase	Carbohydrate metabolism|Nucleotide metabolism
45	16,89	42	valine--tRNA ligase	Translation
46	16,88	42	50S ribosomal protein L5	Translation
53	16,88	42	50S ribosomal protein L6	Translation
42	15,38	40	tRNA pseudouridine(38,39,40) synthase TruA	Translation
73	15,30	38	acetyl-CoA carboxylase biotin carboxylase subunit	Carbohydrate metabolism
43	14,82	37	molecular chaperone DnaK	Translation|Folding, sorting and degradation
72	14,68	41	ribosomal RNA small subunit methyltransferase A	Translation
41	14,19	41	serine--tRNA ligase	Translation
76	13,87	28	preprotein translocase subunit SecY	Membrane transport
67	13,18	41	glutamine--fructose-6-phosphate aminotransferase	Carbohydrate metabolism|Amino acid metabolism|Enzyme families
66	13,00	42	30S ribosomal protein S12	Translation
55	12,42	42	cysteine--tRNA ligase	Translation
48	12,00	42	50S ribosomal protein L11	Translation
103	11,88	28	50S ribosomal protein L3	Translation
174	11,83	35	excinuclease ABC subunit A	Replication and repair
161	10,50	30	Holliday junction DNA helicase RuvB	Replication and repair
70	10,17	27	cell division protein FtsW	Replication and repair
57	9,99	28	30S ribosomal protein S18	Translation
13	9,99	41	ornithine carbamoyltransferase subunit F	Amino acid metabolism
82	9,00	28	elongation factor P	Translation
38	8,99	35	threonine--tRNA ligase	Translation
113	7,88	28	50S ribosomal protein L18	Translation
62	7,15	38	adenylate kinase / nucleoside-diphosphate kinase	Nucleotide metabolism|Metabolism of cofactors and vitamins|Transport and catabolism
33	6,96	28	tyrosine--tRNA ligase	Translation
35	5,91	36	bifunctional 5,10-methylene-tetrahydrofolate dehydrogenase	Energy metabolism|Metabolism of cofactors and vitamins
11	5,88	28	alanine racemase biosynthetic	Drug resistance
86	2,06	37	peptide-methionine (S)-S-oxide reductase	Genetic information processing
68	1,64	38	reactive intermediate/imine deaminase	Metabolism
64	0,98	26	ABC superfamily ATP-binding cassette transporter ABC protein	Membrane transport
59	0,63	22	methionine ABC transporter permease	Membrane transport

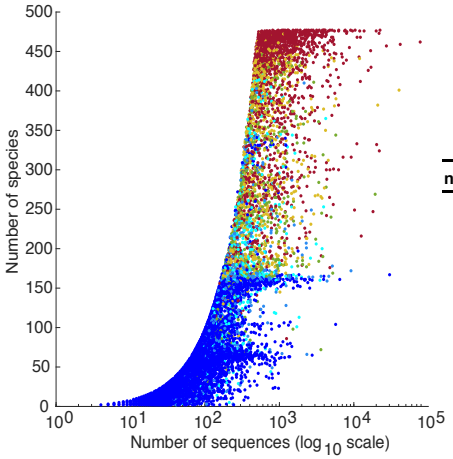
Supplemental Figure 4: Distribution of prokaryote-eukaryote genes across 5,655 prokaryotic genomes. The black boxes indicate if the cluster was determined as widely distributed. The verticality values (extracted from Nagies *et al.*, 2020), number of phyla, NCBI annotation, and KEGG annotation (category B) are indicated as well as the cluster numbers.

a.



	min.	max.	median	mean	std.
n. sequences	4	299	29	38.31	35.95

b.



	min.	max.	median	mean	std.
n. sequences	4	77825	60	205.20	868.89

Supplemental Figure 5. Eukaryotic gene-family size and summary statistics. The plots show the distribution of number of sequences against the number of species across **a)** 24,571 multi-copy gene families clustered with MCL using 150 eukaryotic genomes (data underlying this study); and **b)** 30,063 multi-copy gene families for 477 eukaryotic genomes, assembled on the basis of KOGs (data from eggNOG database version 5 (Huertas-Cepas et al. 2019)). The results show that our stringent clustering cutoffs generate gene-families with number of sequences that are amenable to phylogenetic reconstructions and, at the same time, retaining a high number of paralogous sequences (51% of all eukaryotic sequences distribute in multi-copy families). More liberal approaches based on KOG categories render much larger gene families. In the largest eggNOG family, the number of sequences exceed the number of species by two orders of magnitude. Large gene families of divergent genes are likely to cause errors in downstream phylogenetic analyses (Jeffroy et al. 2006) and, as such, were avoided in this study.

References

Huerta-Cepas, J. *et al.* eggNOG 5.0: a hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses, *Nucleic Acids Res.*, **47**, 309–314 (2019).

Jeffroy O, Brinkmann H, Delsuc F, Philippe H. Phylogenomics: the beginning of incongruence?, *Trends Genet.*, **22**, 225–231 (2006).

Evidence for a Syncytial Origin of Eukaryotes from Ancestral State Reconstruction

Josip Skejo [1,2], Sriram G. Garg [1,*], Sven B. Gould[1], Michael Hendriksen[1], Fernando D. K. Tria [1], Nico Bremer[1], Damjan Franjević[2], Neil W. Blackstone[3], and William F. Martin[1]

[1]Institute for Molecular Evolution, Faculty for mathematics and natural sciences, Heinrich-Heine University Düsseldorf, Düsseldorf, Germany
[2]Faculty of Science, Division of Zoology, Department of Biology, University of Zagreb, Evolution Lab, Zagreb, Croatia
[3]Department of Biological Sciences, Northern Illinois University, DeKalb, IL, USA

*Corresponding author: E-mail: gargs@hhu.de.

Accepted: 28 April 2021

Abstract

Modern accounts of eukaryogenesis entail an endosymbiotic encounter between an archaeal host and a proteobacterial endosymbiont, with subsequent evolution giving rise to a unicell possessing a single nucleus and mitochondria. The mononucleate state of the last eukaryotic common ancestor (LECA) is seldom, if ever, questioned, even though cells harboring multiple (syncytia, coenocytes, and polykaryons) are surprisingly common across eukaryotic supergroups. Here, we present a survey of multinucleated forms. Ancestral character state reconstruction for representatives of 106 eukaryotic taxa using 16 different possible roots and supergroup sister relationships, indicate that LECA, in addition to being mitochondriate, sexual, and meiotic, was multinucleate. LECA exhibited closed mitosis, which is the rule for modern syncytial forms, shedding light on the mechanics of its chromosome segregation. A simple mathematical model shows that within LECA's multinucleate cytosol, relationships among mitochondria and nuclei were neither one-to-one, nor one-to-many, but many-to-many, placing mitonuclear interactions and cytonuclear compatibility at the evolutionary base of eukaryotic cell origin. Within a syncytium, individual nuclei and individual mitochondria function as the initial lower-level evolutionary units of selection, as opposed to individual cells, during eukaryogenesis. Nuclei within a syncytium rescue each other's lethal mutations, thereby postponing selection for viable nuclei and cytonuclear compatibility to the generation of spores, buffering transitional bottlenecks at eukaryogenesis. The prokaryote-to-eukaryote transition is traditionally thought to have left no intermediates, yet if eukaryogenesis proceeded *via* a syncytial common ancestor, intermediate forms have persisted to the present throughout the eukaryotic tree as syncytia but have so far gone unrecognized.

Key words: syncytium, coenocyte, meiosis, mitosis, eukaryogenesis, endosymbiosis, units of selection.

Significance Statement

The transition of prokaryotes to eukaryotes involved endosymbiosis and a dramatic increase in intracellular cell complexity. While most theories on eukaryogenesis consider and illustrate the last eukaryotic common ancestor (LECA) as a mononucleated, sexual, flagellated population of cells, the origin of coordinated nuclear and organellar division coupled to the cell-cycle is rarely discussed. Using ancestral state reconstructions, we show that LECA most likely included a multinucleated stage which also allowed for conflict mediation between mitochondrial and nuclear genomes brought about by endosymbiotic gene transfer. The near-universal presence of the syncytial life stage across all major eukaryotic groups suggests that a multinucleated LECA is a viable intermediate that permitted intracellular experimentation and evolution of the complex eukaryotic processes we observe today.

Introduction

With more than 2 million described species, eukaryotes are morphologically the most diverse domain of life (Archibald et al. 2017; Adl et al. 2019), inhabiting a wide range of ecological habitats (López-García et al. 2007; Mora et al. 2011; Geisen et al. 2017). Eukaryotic cells are vastly more complex than prokaryotic cells as evident by their endomembrane system (Gould et al. 2016). They appear about 2 billion years later in the fossil record than prokaryotes do (Javaux et al. 2001; Javaux and Lepot 2018). There is a consensus among specialists that eukaryotes arose from prokaryotes, but the issue of how they arose from prokaryotes is intensely debated. All current theories for the origin of eukaryotes entail in some manner the concept of symbiogenesis (Mereschkowsky 1910; english translation in Kowallik and Martin 2021) because mitochondria trace to before the last eukaryote common ancestor LECA (Embley and Martin 2006; Tria et al. 2021) and there is no tenable way to explain the structure, DNA, and bioenergetic properties of mitochondria (and chloroplasts) without their endosymbiotic origin. The differences among current theories for eukaryote origin (reviewed in Martin et al. 2015; López-García and Moreira 2015; Dacks et al. 2016) mainly concern assumptions about the biological nature and cellular complexity of the host that acquired the mitochondrion.

In symbiogenic theories, the host is assumed to be a typical archaeon in terms of its cellular complexity, with the origin of mitochondria precipitating genetic, cell biological and bioenergetic changes within the host-symbiont consortium that ultimately led to LECA (Martin and Müller 1998; Lane and Martin 2012; Gould et al. 2016; Imachi et al. 2020). In gradualist theories, the host is assumed to be a descendant of the archaeal lineage, one that had however passed the threshold from prokaryotic to eukaryotic cell complexity by evolutionary mechanisms other than symbiosis, thereby bridging the gap between prokaryotic and eukaryotic complexity (Martijn and Ettema 2013; Spang et al. 2015) before the origin of mitochondria, which therefore had little impact on eukaryote complexity. In hybrid theories, the prokaryote to eukaryote transition involved one or more additional symbioses that preceded the origin of mitochondria, such as flagella (Sagan 1967), peroxisomes (de Duve 1969), the nucleus (López-García and Moreira 2020), or the ER (Gupta and Golding, 1996), or was precipitated by lateral gene transfer (LGT) to the host lineage, such that many hallmark traits of eukaryotes stem from genes that were invented in foreign lineages and donated to LECA via LGT (Pittis and Gabaldón 2016; Vosseberg et al. 2021) although the methods underpinning such claims have been called into question (Martin et al., 2017a; Tria et al. 2021; Nagies et al., 2020). Gradualist and hybrid theories typically posit an origin of phagotrophic feeding within the archaeal host lineage before the origin of mitochondria (Doolittle, 1998; Spang et al., 2015; Zaremba-Niedzwiedzka et al., 2017; Vosseberg et al. 2021), which is however a deeply problematic proposition from the physiological standpoint (Martin et al. 2017b) and at odds with evidence from the microfossil record indicating a late origin of phagocytosis (Mills, 2020). Eukaryotes are unquestionably genetic chimeras, with the majority of eukaryotic genes stemming from bacteria rather than archaea (Brueckner and Martin 2020), wherein the bacterial genes in eukaryotes trace to LECA, not to lineage-specific acquisitions during eukaryotic evolution (Nagies et al., 2020).

Despite their diversity and differing underlying premises, theories for eukaryote origin uniformly entail the assumption, usually implicit, that LECA was unicellular and mononucleate (Gould and Dring 1979; Cavalier-Smith 1987; Lake and Rivera 1994; Gupta and Golding 1996; Horiike et al., 2004; Imachi et al., 2020; Martijn and Ettema, 2013; Martin et al., 2015), an assumption that has almost never been called into question (Garg and Martin 2016). The uniformity of thought on the mononucleate nature of LECA is so pervasive that it is taken as a given, that is, it is rarely, if ever, even mentioned as an assumption. More tellingly, theories for eukaryote origin, if they are illustrated with a schematic diagram at all, invariably convey an image of LECA as a mononucleate cell. Such images are often symbolic in nature, depicting traits as opposed to living cells, but at the same time, they influence the way we conceptualize the problem of eukaryote origin. Models for eukaryogenesis that involve mitochondria in a mechanistic role usually entail one-to-one relationships or many-to-one relationships (Lane and Martin 2012) between mitochondria and the nucleus, whereby the nature of LECA's nuclear dynamics, heterogeneity among nuclei in LECA, its coordination of nuclear division with cell division, its cell cycle (meiotic vs. mitotic) and the evolutionary sequence linking organelle division, nuclear division, and cell division are seldom discussed (Cavalier-Smith 2010; Garg and Martin 2016).

Why is the possibility of a multinucleated state for LECA of interest? The main evolutionary benefit that a multinucleated state would confer upon LECA is evident: Gene mutations or even severe chromosome mutations, including aneuploidies that would otherwise be lethal in a mononucleated cell could be complemented by mRNA from other nuclei in the same cytosol, permitting the survival of the (multinucleated) individual as a collection of heterogeneous nuclei, a stable starting point from which the myriad differences between prokaryotic and eukaryotic chromosome segregation and handling across cell divisions could evolve (Garg and Martin 2016). In this way, the multinucleated state would buffer the transition from prokaryotic to eukaryotic chromosome division and furthermore decouple it from the evolutionary hurdle of surmounting the transition from prokaryotic to eukaryotic cell division as well as prokaryotic to eukaryotic chromatin organization during the cell cycle (Brunk and Martin 2019).

The occurrence of multinucleated taxa has been reported in members of all eukaryotic supergroups and in numerous

higher taxa, some ancient and some derived (Archibald et al. 2017; Adl et al. 2019; see supplementary table 3, Supplementary Material online). Well-known examples of multinucleated forms occur within the amoebozoan supergroup: the myxomycetes (myxogastrid amoebae), protosporangiids, dictyostelids, vampyrellids, and schizoplasmodids (fig. 1). Fungi are perhaps the most common coenocytes on Earth, wherein most of the classes and orders have multinucleated representatives, with unicellular forms being generally rare and often secondarily derived (Kiss et al. 2019). Besides fungi, within opisthokonts, nuclearid amoebae (Dirren and Posch 2016) and ichthyosporeans are also multinucleated, and syncytia are very well known among animals, for example, the body of hexactinellid sponges (Leys 2003), the muscles of all the other animals, and the larvae of holometabolous insects including *Drosophila*. Moreover, it has long been proposed that the common ancestor of Metazoa could have been multinucleated (Hadži 1953). Within Rhizaria, the deepest branch in SAR, there are numerous examples of multinucleated representatives (the most remarkable being Xenophyophorea). Furthermore, Opalinata and Apicomplexa have multinucleated forms as part of their life cycles as well (Archibald et al. 2017; Adl et al. 2019). Not only are syncytia found among heterotrophic eukaryotes but there are also numerous examples of multinucleate algae, both red (*Florideophyceae*) and green (*Ulvophyceae*), as well as various multinucleated tissues in land plants (Niklas et al. 2013). Multinucleated forms also occur among eukaryotes with secondary plastids such as in *Chlorarachniophyceae*, *Phaeophyceae* and *Xanthophyceae* (Niklas et al. 2013). The distribution and evolution of multinucleate tissues among eukaryotes with plastids reveal a great variety of form across 60 archaeplastid families and five diverse algal lineages (Niklas et al. 2013).

Some researchers distinguish between the terms syncytium and coenocyte based on the mechanism underlying the multinucleated state, with syncytia arising from cell fusions and coenocytes arising from chromosome segregation and nuclear divisions, without cytokinesis (Daubenmire 1936). Both lead to a multinucleated state and they are not mutually exclusive. We use the term multinucleated to describe the condition of having more than two (usually four or more) nuclei in the same cell without regard to the mechanism that gave rise to that state. Standard mitotic and meiotic intermediates are, obviously, not scored here as multinucleated states here, as this would trivialize the trait, making it as universal as the presence of nuclei themselves. The images in figure 1 convey an impression of a multinucleated state in the sense intended in this article.

The foregoing observations lead to the question of how far back in eukaryote evolution the syncytial state can be traced. Are multinucleated forms across all eukaryotic supergroups the result of convergence or do they reflect an ancestral state? Here, we explore the presence of multinucleated forms across the breadth of eukaryotic diversity, the likelihood of a multinucleated syncytial LECA using ancestral state reconstruction and the consequences for LECA's lifestyle.

Results

In order to capture the entire diversity of eukaryotes we generated an exhaustive list of 106 eukaryotic taxa (supplementary table 1, Supplementary Material online) including a wide array of organisms with sequenced relatives (see Materials and Methods). Among the 106 taxa chosen only 45 harbor sequenced relatives, highlighting the need for more sequencing of eukaryotic lineages. While there are recent concerted efforts to increase the diversity of sequenced genomes, they still fall short in capturing the immense phenotypic variation that sets the eukaryotes apart from the physiologically diverse prokaryotes. Nevertheless, a sufficient sample of taxa has been studied through microscopy to enable us to tabulate the presence of various eukaryotic traits from the literature, this substantial information is summarized in supplementary tables 2, 3 and 8, Supplementary Material online, including the presence of a multinucleated form in their life cycle. As mentioned in the Introduction section, the multinucleated state is usually reached by one of the two routes, namely through lack of cell division following nuclear division, leading to cells typically designated as coenocytes, and the fusion of mononucleated cells, leading to cells typically designated as syncytia. While the ontogenetic difference between the two states is distinct, in the absence of careful cell biological and cell cycle studies, which are lacking for many of the taxa examined here, it is not possible to accurately code the two form separately and hence are considered here together as being multinucleated. Note that our evolutionary investigation concerns the properties of multinucleated cells and the interactions of nuclei and mitochondria therein, irrespective of the process that generated the multinucleated state.

A cladogram for eukaryotes was generated based on extensive literature (Archibald et al. 2017; Cavalier-Smith 2018; Adl et al. 2019; Kiss et al. 2019, see supplementary tables 2 and 3, Supplementary Material online for complete list) and further refined by allowing for different configurations of polytomies and various accepted positions for the root (altogether 16), resulting in 30 different topologies for the eukaryotes (supplementary table 4, Supplementary Material online). While the tree shown in figure 2 is, among currently available alternatives, the least controversial and possibly most robust tree for those lineages of eukaryotes studied here, all three different topologies were used for ancestral state reconstructions (supplementary table 4, Supplementary Material online).

As control data sets for ancestral state reconstruction, we included several traits that are already annotated for many lineages across the eukaryotic domain. In addition to having mitochondria, the first eukaryote was sexual and had meiotic recombination (Speijer et al. 2015; Fu et al. 2019; Hofstatter

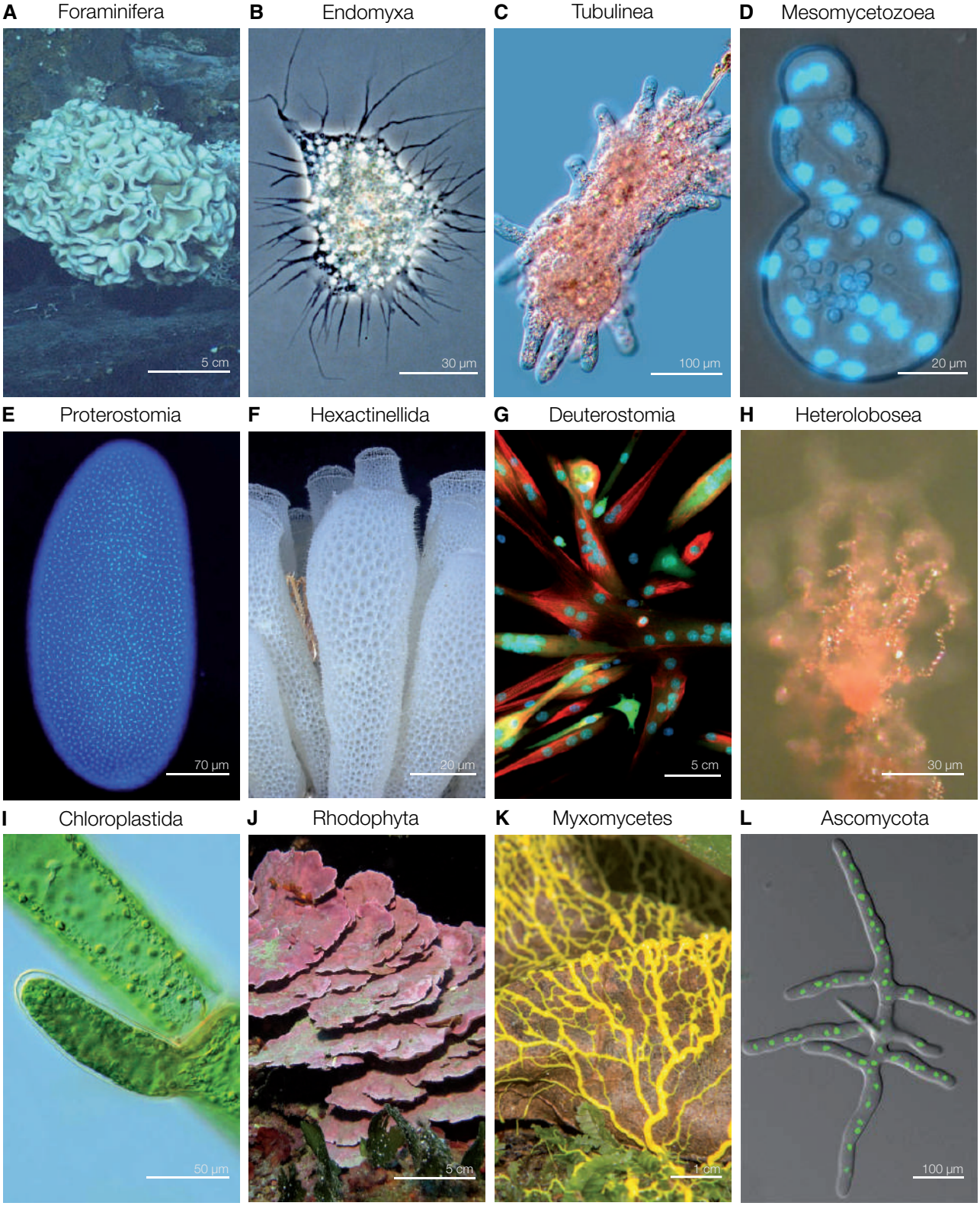


Fig. 1.—Representation of the diversity of the groups harboring multinucleated representatives. (A) Foraminifera: Filosa, a deep sea coenocytic xenophyophore; (B) Endomyxa: *Lateromyxa gallica*, multinucleated predatory amoeba; (C) Tubulinea: *Chaos* sp. multinucleated amoeba; (D) Mesomycetozoea: *Sphaeroforma arctica*, coenocyte with blue nuclei; (E) Protostomia: *Drosophila melanogaster*, multinucleated embryo; (F) Hexactinellida: *Euplectella aspergillum* coenocytic hexactinellid sponge; (G) Deuterostomia: multinucleated mouse muscle cells; (H) Heterolobosea: *Acrasis*

and Lahr 2019). It is known that hydrogenosomes and mitosomes arose from mitochondria via respiratory chain loss and ecological specialization in several independent lineages (Embley and Martin 2006; Müller et al. 2012; Maciszewski and Karnkowska 2019; Gould et al. 2019), that primary plastids arose once (Sánchez-Baracaldo et al. 2017) and that secondary plastids arose several times independently from eukaryotes containing a primary plastid (Maciszewski and Karnkowska 2019; Keeling 2004; Gould et al. 2008). Ancestral state reconstruction should map these traits accordingly.

A general outline of the relationship of prokaryotes to eukaryotes including symbiosis and depicting the number of described species in each group is given in figure 2. Using ancestral state reconstruction, we found that the last eukaryotic common ancestor (LECA) was a sexual, mitochondriate, and heterotrophic organism with closed nuclear division (mitosis) and likely harboring haploid nuclei (figure 3). The method and tree trace sexual reproduction and mitochondria back to the origin of eukaryotic complexity, in agreement with hitherto published studies (Speijer et al. 2015; Hofstatter and Lahr 2019). Lineages with hydrogenosomes, mitosomes and typical mitochondria (fig. 3*A*) represent ecological specializations from a common ancestral organelle (Müller et al. 2012). Consistent with previous reports, sex is recovered as being ubiquitous in the Eukarya domain and meiotic genes are present in all the supergroups in highly conserved manner (Ramesh et al. 2005; Speijer et al. 2015; Hofstatter and Lahr 2019).

The last common ancestor of archaeplastids was the first organism to have a primary plastid and is the first common ancestor of all the (secondary) plastids found in *Euglenida*, *Hacrobia*, and SAR. Despite being widely distributed across the eukaryotic tree, plastids did not trace to LECA with ancestral reconstruction, which serves as a form of internal control (fig. 3*C*). Since LECA did not have a plastid, it could not have been a photosynthetic, autotrophic eukaryote—it was a heterotroph. Primary plastids originated from a cyanobacterium in a symbiogenic event, which likely also involved a freshwater archaeplastid ancestor that was multinucleate in at least part of its life cycle (Sánchez-Baracaldo et al. 2017) (fig. 3*B*). Polyploidy ($>2n$) also originated several times independently. Though polyploid eukaryotes originated numerous times in evolutionarily well-separated groups, LECA most probably had haploid nuclei (fig. 3*A*). Polyploid trophobionts (feeding stages) are rare among eukaryotes. The only polyploid phases in most eukaryotes are the diploid zygote (especially its tetraploid phase before the first division) which ancestrally undergoes meiotic recombination, before the production of four different nuclei (trophic cells, spores, or gametes). In those lineages whose trophobionts are diploid, every somatic nucleus that has DNA replicated before segregation of chromosomes can be regarded as temporarily tetraploid although, as with zygote formation (see Materials and Methods), they were not scored as polyploid. Accordingly, LECA was not polyploid, but because it was meiotic, it harbored some form of karyogamic stage.

While the control traits were reconstructed as expected, the same analysis indicates that LECA was multinucleated and/or had a multinucleated stage during its life-cycle (fig. 3*A* and *B*). It was not a mononucleated protist-like flagellate eukaryote of the type salient to most theories, although it cannot be excluded that some phases of the life cycle might have been mononucleated, protist-like, and flagellated, for example, motile spores. The ancestral reconstruction indicates that the state of LECA might have been multinucleate with nuclei divided by closed mitosis, in which the nuclear envelope remained intact (fig. 3*A*). The ancestral presence of closed mitosis (closed chromosome segregation, not cytokinesis) is significant since chromosome segregation in syncytial forms demands an intact nuclear membrane consistent with an ancestral multinucleated stage. In our analyses, the probability that LECA was multinucleated is as high as the probability that it was sexual and possessed mitochondria (supplementary table 7, Supplementary Material online). For the full detailed results of ancestral character state reconstruction see supplementary table 6, Supplementary Material online.

No matter where we rooted the eukaryotic tree, nor how many unresolved branches we allowed, LECA was always reconstructed as multinucleate. Moreover, and crucially, not only was the ancestor of eukaryotes multinucleated, but the common ancestors of all eukaryote supergroups were also reconstructed as multinucleate as well, except the last common ancestor of Hacrobia (fig. 3*B*). Whether the Hacrobia ancestor was mononucleated or whether the information is missing that lead to inference of a mononucleated Hacrobia ancestor, while all other supergroup ancestors are reconstructed as multinucleated is unresolved. The ancestral reconstruction depicts LECA as multinucleated, a polykaryon whether syncytial or coenocytic, a population of interacting mitochondria and nuclei within the confines of a single cell membrane. Multiple nuclei in the same cytoplasm are not rare

rosea, fruiting body; (I) Chloroplastida: Ulvophyceae: *Cladophora* sp. syphonous thallus; (J)—Rhodophyta: Florideophyceae: *Lithophyllum* sp.; (K)—Myxomycetes: Multinucleated plasmodium of a *Physaraceae* member; (L) Ascomycota: *Eremothecium gossypii*, aseptate hyphae. Photo credits and Creative Commons (CC) sharing domain: A and F. NOAA, public domain; B. Norbert Hülsmann, BY-NC-SA 2.0; C. and I. Proyecto Agua, BY-NC-SA 2.0; D. Multicellgenome lab, BY 2.0; E. Billy Liar, BY-NC-SA 2.0; G. Kevin A. Murach, NIH Image Gallery, BY-CN 2.0; H. Shirley Chio, Biology of Fungi Lab, UC Berkeley, California, BY-SA 3.0; J. Christophe Quintin, BY-NC 2.0; K. André Amaral, distributed under CC BY-NC 4.0; L. Jaspersen Lab, public domain. Scale bar is approximate.

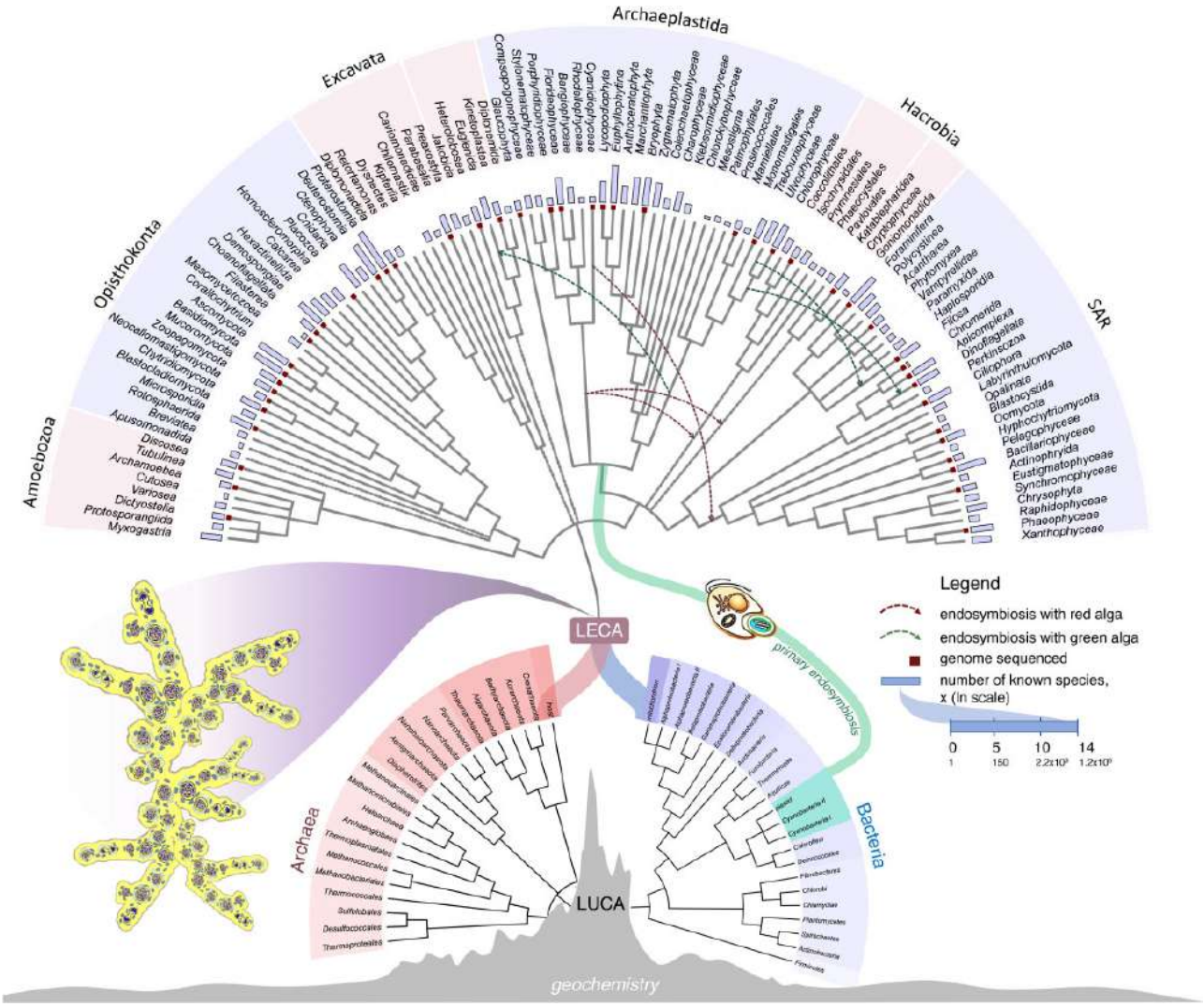


Fig. 2.—Schematic summary of cell evolution. The tree is rooted in physiology and geochemistry, with a nonfree-living Last Universal Common Ancestor (LUCA). Origin of Eukaryotes is depicted as a polyphyletic (symbiogenic) event, where two prokaryotic lineages, an archaeal lineage and an alphaproteobacterial lineage, gave rise to the eukaryotic lineage via LECA—the Last Eukaryotic Common Ancestor. Major prokaryotic groups (within archaea and bacteria) and eukaryotic supergroups are shown, altogether 106 taxa are included in this analysis. Comparison of the number of known species is shown in a logarithmic scale. Squares at the tip of certain branches denote in which groups genomes are sequenced. For reference tree of eukaryotes, see Materials and Methods section. Schematic prokaryotic tree of life was constructed based on literature. The tree was drawn using iTol. LUCA is depicted as arising at a hydrothermal vent, while LECA, which might also have arisen near hydrothermal vents as a geological source of H_2 (15, 23) is depicted as a multinucleate organism in which nuclei divide with their envelopes remaining intact. Primary endosymbiosis with cyanobacteria that gave rise to Archaeplastida is shown. Secondary endosymbiotic events, the multiple origins of secondary plastid, are shown as arrows. A 6-min animated video illustrating the origin of eukaryotes from symbiosis and the role of a syncytial state in the life cycle of LECA can be viewed at (https://www.youtube.com/watch?v=mmh_IpdgWvw&t=2s).

phenomena among eukaryotes. Syncytia and coenocytes are found across most of the higher eukaryotic groups (supplementary table 2, Supplementary Material online). Free-floating nuclei in the cytosol as opposed to being tethered to cell walls imply that in a syncytium they can only divide if nuclear division—and consequently chromosome segregation—is closed wherein the nuclear membrane remains intact throughout mitosis. Open mitosis in a coenocyte would potentially result in the spindle apparatus attaching to chromosomes from different nuclei and segregating them in an aberrant and likely lethal manner. The reconstruction (fig. 3*B*; table 1) suggests that open nuclear division (dissolution of the nuclear membrane at mitosis, as is well known in vertebrates) originated from closed mitosis via semi-open division, in which parts of the nuclear envelope dissolve, as the intermediate state (Boettcher and Barral 2013). Open nuclear division is typical for some mononucleate (both unicellular and multicellular), and most land-inhabiting eukaryotes (fig. 3*A*).

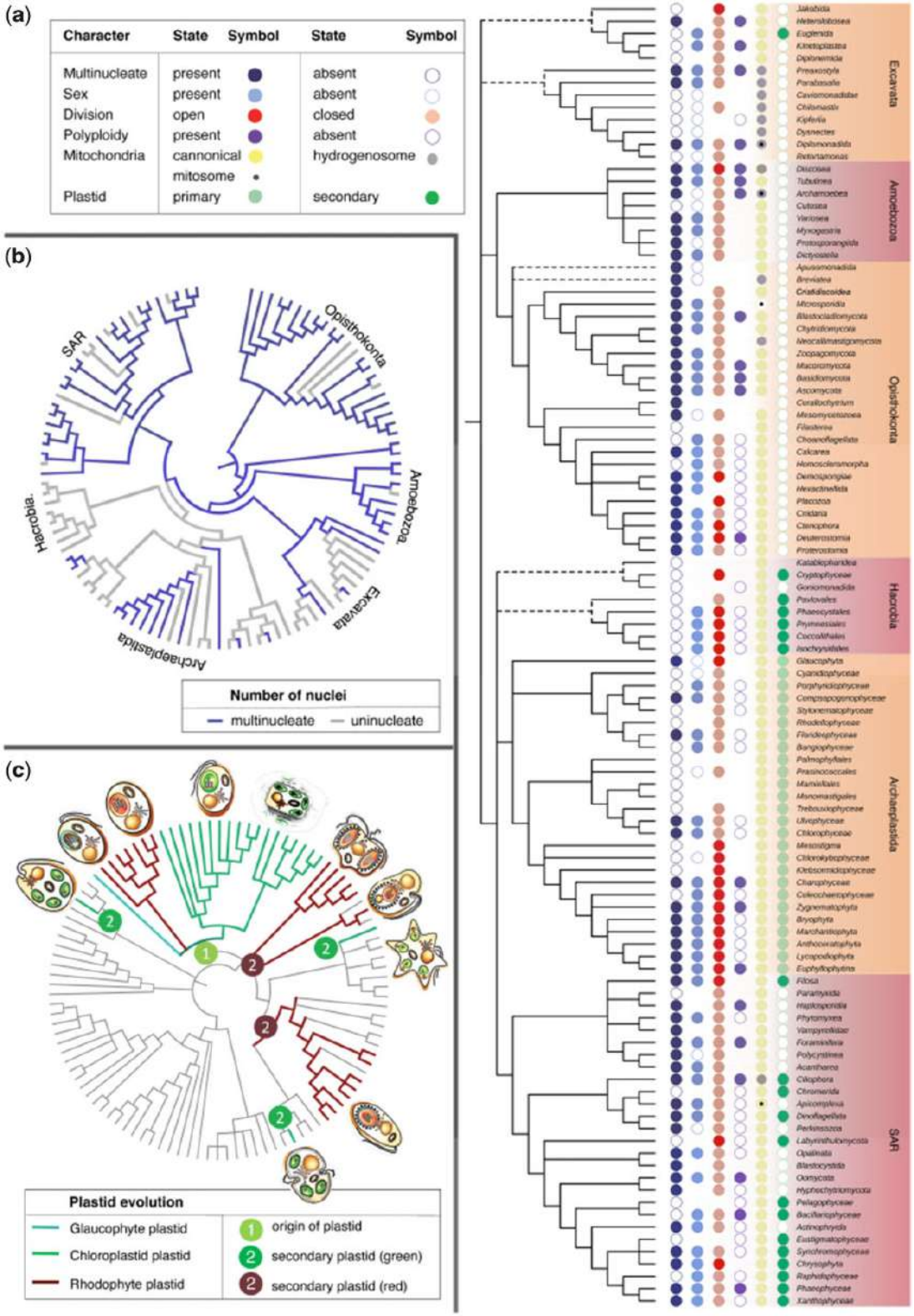


Fig. 3.—Overview of eukaryotic diversity with ancestral state reconstruction. (A) Traits are annotated on the reference eukaryote tree (1. presence of the multinucleated state, 2. sexual reproduction, 3. type of nuclear division, 4. polyploidy, 5. type of mitochondria and 6. type of plastids). (B) Ancestral state reconstruction of the multinucleate state is shown, as well as (C). the plastid evolution Representatives of main photosynthetic eukaryotes are depicted schematically. The data indicate that LECA was a multinucleate sexual heterotroph with closed mitosis.

Table 1
Summary of the results of the ancestral character state reconstruction by maximum parsimony, across 30 different topologies

	% Trees		
	Absent	Present	Ambiguous
Syncytium	0	0.8	0.2
Sex	0	0.867	0.133
Polyploidy	1	0	0
Closed division	0	1	0
Mitochondria	0	0.867	0.133
Plastid	1	0	0

Discussion

The origin of eukaryotes was a unique event from which all the complex life stems. The symbiosis that gave rise to eukaryotes occurred over 1.5 billion years ago (Knoll et al. 2006). While eukaryote origin cannot be forced to occur in the laboratory, endosymbiosis can (Mehta et al. 2018). The contours of eukaryogenesis, intermediate stages, and the sequence of events involved can be addressed via inference from the comparative investigation of modern lineages. The first eukaryote was the result of interactions between archaea and bacteria, two highly divergent cell lineages, that gave rise via interaction and cooperation to a new kind of organism, LECA, with new properties, novel bioenergetics, chimeric chromosomes, a cell cycle, novel genetics, reciprocal recombination, and cellular complexity. Descendants of these symbiotic partners are preserved as bacterial ribosomes in mitochondria and archaeal ribosomes in the eukaryotic cytosol.

LECA had sexual reproduction that included the fusion of haploid nuclei selected for the reproduction (gametes) and the recombination of their genetic material (meiosis). Mitochondria, sex, and multiple nuclei are signatures of LECA's state, with synergistic interactions. Unlike mitochondria, the nucleus has a large, complex genome with little size constraint. The genetic compatibility of nuclei and mitochondria inhabiting the same cytoplasm is crucial for the survival of eukaryotic cells. Internal competition or cytonuclear incompatibility can be lethal (Blackstone and Green 1999; Pesole et al. 2012; Rand and Mossman 2020) or render the organism dysfunctional. Inheritance of mitochondria is often uniparental. The inheritance of the nuclear genome is, however, bi-, tri-, or multi-parental. Uniparental inheritance of mitochondria indicates the existence of strict control on compatibility. Meiotic recombination, ancestrally during the zygote phase, is a compatibility checkpoint. At the onset of eukaryote evolution, the compatibility of mitochondria with newly arisen nuclei was essential. In mononucleate cells, only compatible combinations survived natural selection. In syncytia, many-to-many interactions among mitochondria and nuclei buffered compatibility within the environmental confines of a single cytoplasm. Spores spawned from a syncytial LECA presented a powerful bottleneck of selection for cytonuclear compatibility (Garg and Martin 2016).

An intriguing aspect of the multinucleated state for LECA concerns the transition from prokaryotic to eukaryotic chromosome segregation. In prokaryotes, chromosome segregation is linked to cell division via chromosome attachment to the cell wall. In eukaryotes, microtubule-dependent segregation of condensed chromosomes and cell division (cytokinesis) are neither physically nor mechanistically linked, though often temporally apposed. That is, chromosomes can, and often do, replicate and segregate in nondividing cells without the formation of spindles for the division of the nucleus itself (Geitler 1953), processes that were termed *Amitose* in the older literature (Strasburger 1908). If the origin of nuclear division (replication followed by segregation) preceded the origin of cell division at eukaryote origin (the converse could hardly be true), the resulting syncytium need not have possessed well-regulated chromosome segregation at the outset. It could have generated nuclei with aberrant chromosome numbers or aneuploid haploids. Such defective nuclei would be lethal for a mononucleate cell, but not in a syncytium, because even highly defective nuclei could complement each other freely via mRNA in the cytosol. The multinucleate state would thus buffer virtually all deleterious effects of nuclei arising as products of incorrect chromosome partitioning during a closed protomitosis at the origin of eukaryote chromosome segregation. This would have kept the syncytium as a unit of vegetative proliferation alive, while harboring nuclei with very different chromosome sets, nuclei that kept each other viable within the syncytium through complementation via mRNA in the cytosol. This involvement of ribosomes, whose synthesis requires massive rRNA gene expression, for complementation would explain why the nucleus: cytoplasm volume ratio (*Kern-Plasmarelation*) tends to approach a roughly constant value (Klieneberger 1917) of 1:10 even in syncytial cells (Sitte et al. 1991). As Strasburger (1908) put it: "In the Characeae, amitotic nuclear division in internodial cells is not a degenerate process, rather it is a means to amplify certain components of nuclear substance in relationship to the increase of cytoplasmic mass" (p. 40, translation by the authors).

Physical fusion of nuclei, a primitive and unregulated forerunner of karyogamy (present in LECA because LECA had sex), would generate new combinations of chromosomes at the same time as genes were being transferred from mitochondria to the nuclei (Lane and Martin 2012; Garg and Martin 2016). That generated a heterogeneous population of nuclei interreacting with a heterogeneous population of mitochondria, within the same syncytium. A syncytium could also become physically severed, generating segments or fragments that, provided means of sealing off ends, could have generated descendant progeny (as diaspores) without the requirement for regulated cell division. Syncytial fragments provided a mechanism for propagating populations of nuclei and mitochondria. But the main evolutionary hurdle to be crossed

was evolution of regulated, symmetric chromosome segregation that took into account the nutritional state of the cell (Brunk and Martin 2019) en route to a cell cycle—the backbone of eukaryotic cell biology.

Within a syncytium, both nuclei and mitochondria were units of selection and units of evolution. They were the intermediate state in the prokaryote to eukaryote transition. They coexisted within the same cytosol. Nuclei became heritable collections of genes able to influence their immediately surrounding cytosol, and able to interact with each other and with mitochondria via exported mRNA. Multinucleated cells are ubiquitous among the eukaryotes, both living (figs. 1 and 3A) and fossil, such as a recently reported 1-billion-year-old coenocytic green alga (Tang et al. 2020).

Conflict and Co-operation in a Syncytial LECA

Mitonuclear compatibility is important and is proportional to cell fitness (Rand and Mossman 2020). To compare the relative fitness of a mononucleated cell (monokaryon) and a multinucleated cell (polykaryon), one can consider the difference between the probability of survival for a population of unicellular mononucleate eukaryotes versus that for a single syncytium. For monokaryons, the probability of survival of the population is dependent on the individual survival probabilities which in turn depend on the fitness of the respective mitonuclear pair. However, in the case of a syncytium since the mitochondria and nuclei coexist in one cell the survival probability depends on the cumulative fitness of all possible combinations of mitonuclear pairs. This in turn allows the syncytium to behave similar to a population while allowing selection to resolve internal mito-nuclear conflicts independently. This is schematically shown in figure 4 and mathematically described in supplementary information 11, Supplementary Material online. A syncytium behaves as more like a population of nuclei and mitochondria than as an individual cell. Thus, the syncytium has a higher chance of survival than a population of monokaryons. Of course, there are ancient lineages of eukaryotes harboring mononucleate forms, including the excavates. However, a multinucleated LECA explains why modern eukaryote diversity is more readily derived from a syncytial ancestor than from a population of mononucleate unicellular ancestors (monokaryons). A population of monokaryons, especially that of haploid monokaryons, is not likely to accumulate genetic diversity. A syncytium on the other hand, easily accumulates genetic diversity within one cytosol, as nuclei with advantageous alleles complement deficiencies of other nuclei, and karyogamy, of which a meiotic LECA was capable, within a syncytium can generate novel chromosome combinations (fig. 4).

Evolutionary transitions in individuality involve cooperation and conflict (Buss 1987; Maynard Smith and Szathmáry 1996; Michod 1999). Without mechanisms for conflict mediation, cooperation cannot survive (Nowak 2006) and the higher-level unit cannot emerge (Radzvilavicius and Blackstone 2018). In evolution, the population structure has always been recognized as one of the most general mechanisms favoring cooperation. Even if selection favors non-cooperating defectors, as is typically the case, cooperation might still evolve in a structured population. Consider a population made up of individuals (the lower level) divided into groups (the higher level). While defectors are favored at the lower level, cooperators are favored at the higher. If a population was one large group, the selection at the higher level is weak, and defectors prevail. In a population with many small groups, however, the selection is potentiated at the higher. Groups of cooperators can form by chance and outcompete groups of defectors (Szathmáry and Demeter 1987). Thus, larger groups (e.g., a syncytial LECA) invite more conflict, while smaller groups (particularly sexually produced gametes) entail less. With larger cell sizes, stochastic processes may hence have been less important in mediating evolutionary conflict.

Origins of Flagellated Eukaryotes

In comparison to prokaryotes, the eukaryotic cell cycle is as unique as the processes behind mitosis and the physical separation of the newly emerging cell (cytokinesis). While a few homologous proteins are shared between archaeal binary fission and eukaryotic cytokinesis (Lindås et al. 2008), the mechanism of chromosome segregation through a centrosome-organized microtubular system and the subsequent actin-based cell constriction is not conserved across the prokaryote–eukaryote divide. The mechanism of eukaryotic chromosome segregation, like other eukaryote-specific traits, evolved de novo during the endosymbiotic integration of a bacterial partner within an archaeal cytosol en route to LECA. In a syncytium, chromosome segregation likely involves molecular selforganization of a chromosome separating machinery that requires no anchoring points at the plasma membrane.

Closed mitosis, in which the nuclear envelope remains largely intact, is considered ancestral to open mitosis (Cavalier-Smith 2010), consistent with our own results (fig. 3). All variants of mitosis share a microtubule-based network, which can be bundled or loose in a star-like manner, that reach out for the chromosomes and attach at the kinetochore which was present in LECA (Tromer et al. 2019). Centrosomes are however, not essential for chromosome separation (Heald et al. 1996). Crucially, eukaryotic chromosomes are separated largely by pushing forces along microtubules, in which the eukaryote-specific kinesin family of proteins play an essential role (Shimamoto et al. 2015). These mechanisms fit seamlessly with the biology of a syncytial cell, as mitosis of individual nuclei can occur independently of localized plasma membrane fixation points.

Consequently, the origin of mononucleated, flagellated protists can be viewed from a novel perspective. Images of the closest living relative of the archaeal host cell and a

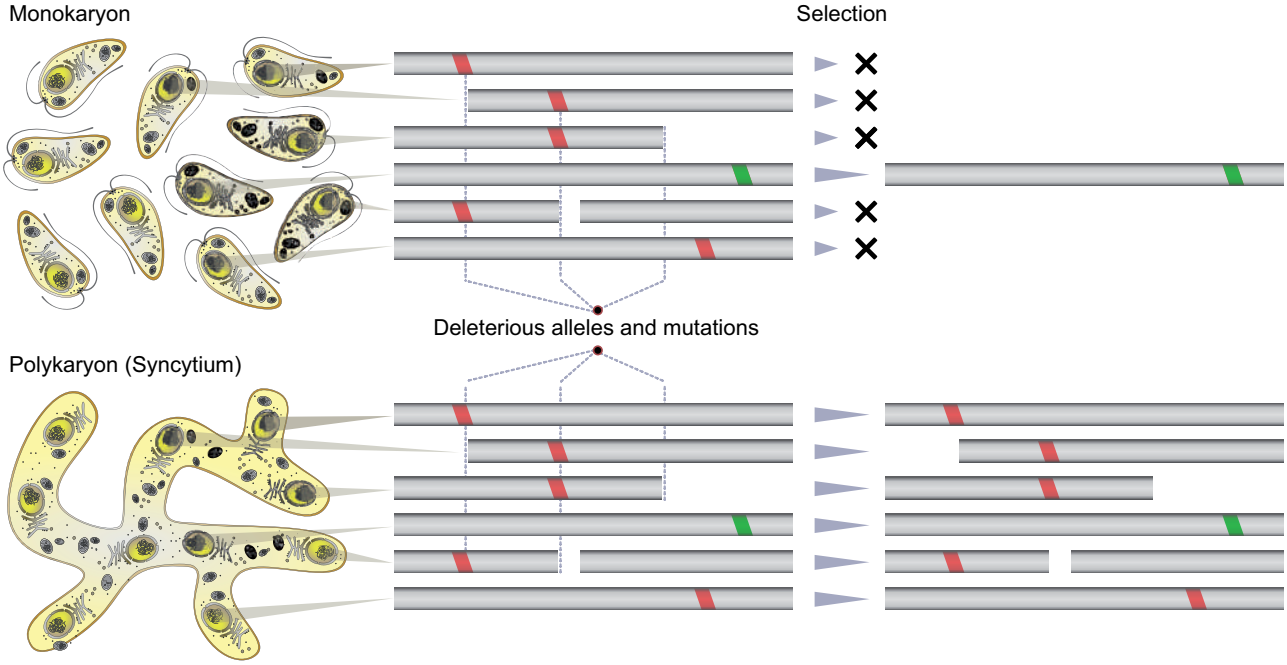


Fig. 4.—Syncytia buffer chromosome defects, unlike monokaryons. Schematic representation of a population of unicellular protists and a syncytial cell. Genomes of each nucleus are schematically shown as grey lines, deficient alleles as red rectangles whereas the beneficial allele is shown in green. If the same evolutionary constraints are applied, monokaryons' population is more likely to go extinct than a syncytium, as nuclei from different cells can neither cover each other's defects nor buffer mitonuclear incompatibilities, while in syncytium they can.

bacterial partner depict two sessile, nonmotile partners (Imachi et al. 2020), the syncytial LECA we propose was sessile, too. The microtubule organizing center (MTOC), or basal body, and the ability to form flagella was present in LECA. This trait diversified among eukaryotic supergroups and underwent recurrent loss (Yubuki and Leander 2013). Eukaryotic flagella are directly connected to basal bodies, or they form in a centriole-dependent manner de novo (Schrøder et al. 2011). The flagella pore complex shares a number of proteins with the nuclear pore complex (Dishinger et al. 2010; Kee et al. 2012; Gould et al. 2016). We suggest that the flagellum evolved on the basis of a (duplicated) centrosome-derived structure that subtended a region of the plasma membrane. Mononucleate, flagellated spores could have thus emerged from the syncytium with the actin cytoskeleton supporting final scission (Heidstra 2007).

Only spores containing viable mitonuclear interactions and capable of flagellar motion would have had the properties of motile gametes, provided that they were able to fuse with others of their kind, which is possible given the tendency of archaea themselves to fuse (Lange et al. 2011; Garg and Martin 2016; Shalev et al. 2017). Such spores would present motile units of selection. The nucleus of many flagellated protists is located in close proximity to the basal body, if not connected to it, as in numerous Archamoebea, Chytridiomycota, *Olpidium*, Pelagophyceae, Bacillariophyceae, Rhizaria and others (reviewed in ref. 2). It is possible that such gamete like cells became the founders of eukaryotic supergroups, all of which contain flagellated representatives that can generate syncytia (fig. 3*A*). We have no suggestion for the physical size of LECA as a syncytium, although we do suggest that it was a marine sediment dweller (Martin and Müller 1998), where anaerobic syntrophy is essential to symbiotic interactions (Imachi et al. 2020). The hyphae of modern fungal individuals can cover areas of square miles (Anderson et al. 2018). LECA could have been a large non-dividing multinucleate unicell that spawned supergroups through the extrusion of mitochondriate flagellated spores. A 6-min animated video illustrating the origin of eukaryotes from symbiosis and the role of a syncytial state in the life cycle of LECA can be viewed at (https://www.youtube.com/watch?v=mmh_IpdgWvw&t=2s)

Conclusion

Unlike prokaryotes, eukaryotes have complex systems of intracellular membrane flux and possess organelles. They are in terms of morphology the most diverse domain of life and originated via the origin of mitochondria. Eukaryote origin is usually depicted as a narrative of *two-cells-becoming-one*, a *one-on-one-model*, where an archaeon host engulfed a proteobacterial symbiont, with the units of selection being chimeric, mononucleate, free-living cells. Our results however suggest that at eukaryote origin, nuclei, and mitochondria were the units of selection and the units of evolution within

the confines of a syncytial LECA. Ancestral character state reconstruction based on taxon rich sampling spanning all supergroups suggest that LECA was 1) mitochondriate, 2) multinucleate (syncytial, coenocytic), 3) haploid, 4) with closed nuclear division, and 5) with sexual reproduction. It is often stated, also in many papers by the present authors, that the prokaryote to eukaryote transition left no intermediate forms. However, if our current thoughts are roughly on target, syncytia are in fact the intermediate state in the prokaryote to eukaryote transition, though hitherto unrecognized as such. In that light, the syncytia present throughout all eukaryote supergroups may harbor previously unrecognized forms of evidence about eukaryote origin and the prokaryote to eukaryote transition.

Materials and Methods

Selection of taxa

Based on an inspection of the literature (supplementary table 4, Supplementary Material online), a taxon-rich (Katz and Grant 2014) eukaryotic dataset comprising 106 higher taxa was constructed (supplementary table 1, Supplementary Material online). Representatives of six eukaryote supergroups are included. We employ the nomenclature of eukaryote supergroups as recently defined: *Amoebozoa, Archaeplastida, Excavata, Hacrobia, Opisthokonta*, and *SAR*, although for clarity, we have retained the more familiar term Opisthokonta instead of Obazoa here. The set consists mostly of higher categories, but in some cases, families and genera were included (table 2).

Reference tree construction

Eukarya includes six supergroups—*Archaeplastida*, *Amoebozoa*, *Excavata*, *Hacrobia*, *Opisthokonta*, and *SAR*. Cladograms represented in this study are based on published relationships within each eukaryote supergroup (supplementary tables 3 and 4, Supplementary Material online). If we designate relationships as "resolved" it means that we incorporated the corresponding branching pattern for our 106-taxa tree. Branching patterns that were "unresolved" were translated to polytomies.

Tree topology and root

With or without a resolved species tree, the root branch is always informative and the output of phylogenetic analysis can vary depending on the position of the root (Tria et al. 2021). Within the supergroups, we deal mostly with resolved "species trees" (see Reference tree). However, relationships between the supergroups are not completely resolved, so we employed two models in ancestral state reconstruction—one that allows polytomy (unresolved branches), and the other which allows only dichotomies. Because there is no consensus on where the eukaryote root Eukarya lies, a set of reference trees was prepared with a collection of published proposals for the eukaryote root: *Excavata* or within excavates (Cavalier-Smith 2002; He et al. 2014; Tria et al. 2021), *Opisthokonta*, *Fungi* or within *Fungi* (e.g., *Microsporidia*) (Vossbrinck et al. 1987), *Amoebozoa* or within (Stechmann and Cavalier-Smith 2002; Katz and Grant 2015), *Amorphaea* (*Opisthokonta*+*Amoebozoa*) (Derelle et al. 2015). An unrooted set was also prepared. Detailed data underlying all parameters and trees are presented in supplementary table 3, Supplementary Material online.

Table 2
Overview of eukaryote taxa considered in this analysis

	Phyla	Classes	Orders	Families	Genera	Per Supergroup
Amoebozoa	—	7	1	—	—	8
Archaeplastida	7	14	4	—	1	26
Excavata	1	4	3	1	4	13
Hacrobia	—	3	5	—	—	8
Opisthokonta	13	9	1	—	1	24
SAR	6	17	3	1	—	27
per rank	27	54	17	2	6	

Number of phyla, classes, orders, families, and genera are shown, corresponding to each supergroup (*per* supergroup).

Annotation of traits

To address LECA's traits, a data set comprising six characters was assembled: (I) the multinucleate state, (II) sexual, meiotic reproduction, (III) behavior of the nuclear envelope during division, (IV) polyploidy (>2*n*), (V) type of mitochondria, and (VI) presence and type of plastid. All the traits were numerically coded for ancestral state reconstruction (supplementary table 5, Supplementary Material online). For the multinucleated state, the trait was coded as 1 when there was an indication of the multinucleated state present in the whole group or part of the lifecycle of many (>2 genera) members, or if it was present within unresolved groups. The trait was considered ambiguous (0/1) for a group when either there is either a consensus that multinucleate state is present in a single derived species within the group, or when there is evidence for the presence of life cycle stages that closely resemble syncytia-like structures without a clear description in members of a group. Finally, the trait was coded as 0 when there is no indication at all that a multinucleate state exists within the known diversity of a certain taxon. The sources for the information are summarized in the supplementary tables 2 and 4, Supplementary Material online.

Ancestral character state reconstruction

Analyses of the ancestral state reconstruction were performed on the basis of the numerically coded character matrix

(supplementary table 5, Supplementary Material online) using maximum parsimony (ordered and unordered) (supplementary table 6, Supplementary Material online) as implemented in Mesquite 3.6 software (https://www.mesquiteproject.org/). Altogether 106 eukaryote taxa (from genus to phylum level) were selected (see Selection of taxa, supplementary table 1, Supplementary Material online). A set of 30 reference trees was then constructed, with different positions of root and different freedom towards polyphyly (see Setting tree topology and root, supplementary table 3, Supplementary Material online). The character matrix was prepared from literature data for six traits: the presence of multinucleated state, presence of polyploids, presence of sex/meiosis, behavior of nuclear envelope during division, type of mitochondria, and type of plastid (supplementary tables 2 and 5, Supplementary Material online). In some groups, certain members of the group exhibit one trait, while others exhibit the other. In cases like this, both traits were coded for that group (0/1 or 1/2).

Acknowledgments

This research was supported and funded by the Moore–Simons Project on the Origin of the Eukaryotic Cell GBMF9743 to W.F.M. and S.G.G. W.F.M. also would like to thank the ERC (666053) and the VW Foundation (93046 and 96742) for funding. S.B.G would like to thank VolkswagenStiftung (Life) for funding. We thank Prof. Klaus V. Kowallik for helpful discussions.

Author Contributions

All authors read the manuscript contributed to the final version of the manuscript. There are no conflicts of interests between the co-authors.

Data availability

All data associated with the manuscript is provided in the supplementary information.

References

Adl SM, et al. 2019. Revisions to the classification, nomenclature, and diversity of eukaryotes. J Eukaryot Microbiol. 66(1):4–119.
Anderson JB, et al. 2018. Clonal evolution and genome stability in a 2500-year-old fungal individual. Proc Biol Sci. 285(1893):20182233.
Archibald JM, Simpson AGB, Slamovits CH, editors. 2017. Handbook of the protists. Springer Nature.
Blackstone NW, Green DR. 1999. The evolution of a mechanism of cell suicide. Bioessays. 21(1):84–88.
Boettcher B, Barral Y. 2013. The cell biology of open and closed mitosis. Nucleus 4(3):160–165.
Brueckner J, Martin WF. 2020. Bacterial genes outnumber archaeal genes in eukaryotic genomes. Genome Biol Evol. 12(4):282–292.
Brunk CF, Martin WF. 2019. The archaeal histone contribution to the origin of eukaryotes. Trends Microbiol. 27(8):703–714.
Buss LW. 1987. The evolution of inidividuality. New Jersey: Princeton University Press.
Cavalier-Smith T. 1987. The origin of eukaryote and archaebacterial cells. Ann N Y Acad Sci. 503:17–54.
Cavalier-Smith T. 2002. The phagotrophic origin of eukaryotes and phylogenetic classification of Protozoa. Int J Syst Evol Micro. 52(2):297–354.
Cavalier-Smith T. 2010. Origin of the cell nucleus, mitosis and sex: roles of intracellular coevolution. Biol Direct. 5:7.
Cavalier-Smith T. 2018. Kingdom Chromista and its eight phyla: a new synthesis emphasising periplastid protein targeting, cytoskeletal and periplastid evolution, and ancient divergences. Protoplasma 255(1):297–357.
Dacks JB, et al. 2016. The changing view of eukaryogenesis – fossils, cells, lineages and how they all come together. J Cell Sci. 129(20):3695–3703.
Daubenmire RF. 1936. The use of the terms coenocyte and syncytium in biology. Science. 84(2189):533–533.
de Duve C. 1969. Evolution of the peroxisome. Ann N Y Acad Sci. 168(2):369–381.
Derelle R, et al. 2015. Bacterial proteins pinpoint a single eukaryotic root. Proc Natl Acad Sci U S A. 112(7):E693–E699.
Dirren S, Posch T. 2016. Promiscuous and specific bacterial symbiont acquisition in the amoeboid genus *Nuclearia* (Opisthokonta). FEMS Microbiol Ecol. 92(8):fiw105.
Dishinger J, et al. 2010. Ciliary entry of the kinesin-2 motor KIF17 is regulated by importin-β2 and RanGTP. Nat Cell Biol. 12(7):703–710.
Doolittle WF. 1998. You are what you eat: a gene transfer ratchet could account for bacterial genes in eukaryotic nuclear genomes. Trends Genet. 14(8):307–311.
Embley TM, Martin WF. 2006. Eukaryotic evolution, changes and challenges. Nature 440(7084):623–630.
Fu C, Coelho MA, David-Palma M, Priest SJ, Heitman J. 2019. Genetic and genomic evolution of sexual reproduction: echoes from LECA to the fungal kingdom. Curr Opin Genet Dev. 58-59:70–75.
Garg SG, Martin WF. 2016. Mitochondria, the cell cycle, and the origin of sex via a syncytial eukaryote common ancestor. Genome Biol Evol. 8(6):1950–1970.
Geisen S, et al. 2017. Soil protistology rebooted: 30 fundamental questions to start with. Soil Biol Biochem. 111:94–103.
Geitler L. 1953. Endomitose und amitotische Polyploidisierung. Protoplasmatologia. Handbuch Der Protoplasmaforschung Bd. VI: Kern- Und Zellteilung. Vol. 6:C. Springer, Wien. p. 89.
Gould GW, Dring GJ. 1979. On a possible relationship between bacterial endospore formation and the origin of eukaryotic cells. J Theor Biol. 81(1):47–53.
Gould SB, Waller RF, McFadden GI. 2008. Plastid evolution. Annu Rev Plant Biol. 59:491–517.
Gould SB, Garg SG, Martin WF. 2016. Bacterial vesicle secretion and the evolutionary origin of the eukaryotic endomembrane system. Trends Microbiol. 24(7):525–534.
Gould SB, et al. 2019. Adaptation to life on land and high oxygen via transition from ferredoxin- to NADH-dependent redox balance. Proc R Soc B. 286(2019149).
Gupta RS, Golding GB. 1996. The origin of the eukaryotic cell. Trends Biochem Sci. 21(5):166–171.
Hadži J. 1953. An attempt to reconstruct the system of animal classification. Syst Zool. 2:145–154.
He D, et al. 2014. An alternative root for the eukaryote tree of life. Curr Biol. 24(4):465–470.,
Heald R, et al. 1996. Self-organization of microtubules into bipolar spindles around artificial chromosomes in *Xenopus* egg extracts. Nature 382(6590):420–425.

Heidstra R. 2007. Asymmetric cell division in plant development. In: Macieira-Coelho A, editor. Asymmetric cell division. Verlag Berlin Heidelberg: Springer. p. 1–37.
Hofstatter PG, Lahr DJ. 2019. All eukaryotes are sexual, unless proven otherwise: many so-called asexuals present meiotic machinery and might be able to have sex. Bioessays 41(6):e1800246.
Horiike T, Hamada K, Miyata D, Shinozawa T. 2004. The origin of eukaryotes is suggested as the symbiosis of *Pyrococcus* into γ-proteobacteria by phylogenetic tree based on gene content. J Mol Evol. 59(5):606–619.
Imachi H, et al. 2020. Isolation of an archaeon at the prokaryote-eukaryote interface. Nature 577(7791):519–525.
Javaux EJ, Knoll AH, Walter MR. 2001. Morphological and ecological complexity in early eukaryotic ecosystems. Nature 412(6842):66–69.
Javaux EJ, Lepot K. 2018. The paleoproterozoic fossil record: implications for the evolution of the biosphere during earth's middle-age. Earth-Sci Rev. 176:68–86.
Kamikawa R, et al. 2014. Gene content evolution in discobid mitochondria deduced from the phylogenetic position and complete mitochondrial genome of *Tsukubamonas globosa*. Genome Biol Evol. 6(2):306–315.
Katz LA, Grant JR. 2015. Taxon-rich phylogenomic analyses resolve the eukaryotic tree of life and reveal the power of subsampling by sites. Syst Biol. 64(3):406–415.
Kee HL, et al. 2012. A size-exclusion permeability barrier and nucleoporins characterize a ciliary pore complex that regulates transport into cilia. Nat Cell Biol. 14(4):431–437.,
Keeling PJ. 2004. Diversity and evolutionary history of plastids and their hosts. Am J Bot. 91(10):1481–1493.
Kiss E, et al. 2019. Comparative genomics reveals the origin of fungal hyphae and multicellularity. Nat Commun. 10(1):13.
Klieneberger E. 1917. Über die Größe und Beschaffenheit der Zellkerne mit besonderer Berücksichtigung der Systematik (Inaug.-Diss. Frankfurt). Beihefte Bot. Zbl. 35 I:219–278.
Knoll AH, Javaux EJ, Hewitt D, Cohen P. 2006. Eukaryotic organisms in Proterozoic oceans. Philos Trans R Soc Lond B Biol Sci. 361(1470):1023–1038.
Kowallik KV, Martin WF. 2021. The dawn of symbiogenesis: annotated English translation of Mereschkowsky's 1910 paper. BioSystems 199:104281.
Lake JA, Rivera MC. 1994. Was the nucleus the first endosymbiont? Proc Natl Acad Sci U S A. 91(8):2880–2881.
Lane N, Martin WF. 2012. The origin of membrane bioenergetics. Cell 151(7):1406–1416.
Lindås AC, et al. 2008. A unique cell division machinery in the Archaea. Proc Natl Acad Sci U S A. 105(48):18942–18946.
Lange C, Zerulla K, Breuert S, Soppa J. 2011. Gene conversion results in the equalization of genome copies in the polyploid haloarchaeon *Haloferax volcanii*. Mol Microbiol. 80(3):666–677.
Leys SP. 2003. The significance of syncytial tissues for the position of the Hexactinellida in the Metazoa. Integr Comp Biol. 43(1):19–27.
López-García P, Moreira D. 1999. Metabolic symbiosis at the origin of eukaryotes. Trends Biochem Sci. 24(3):88–93.
López-García P, Vereshchaka A, Moreira D. 2007. Eukaryotic diversity associated with carbonates and fluid–seawater interface in Lost City hydrothermal field. Environ Microbiol. 9(2):546–554.
López-García P, Moreira D. 2015. Open questions on the origin of eukaryotes. Trends Ecol Evol. 30(11):697–708.
López-García P, Moreira D. 2020. The syntrophy hypothesis for the origin of eukaryotes revisited. Nat Microbiol. 5(5):655–667.
Maciszewski K, Karnkowska A. 2019. Should I stay or should I go? Retention and loss of components in vestigial endosymbiotic organelles. Curr Opin Genet Dev. 58-59:33–39.
Martijn J, Ettema TJG. 2013. From archaeon to eukaryote: the evolutionary dark ages of the eukaryotic cell. Biochem Soc T. 41(1):451–457.
Martin WF, Müller M. 1998. The hydrogen hypothesis for the first eukaryote. Nature 392(6671):37–41.
Martin WF, Garg S, Zimorski V. 2015. Endosymbiotic theories for eukaryote origin. Philos Trans R Soc Lond B Biol Sci. 370(1678):20140330.
Martin WF, et al. 2017a. Late mitochondrial origin is an artifact. Genome Biol. Evol. 9(2):373–379.
Martin WF, Tielens AGM, Mentel M, Garg SG, Gould SB. 2017b. The physiology of phagocytosis in the context of mitochondrial origin. Microbiol Mol Biol Rev. 81(3):17.
Maynard Smith J, Szathmáry E. 1995. The major transitions in evolution. Oxford: Oxford University Press.
Mehta AP, et al. 2018. Engineering yeast endosymbionts as a step toward the evolution of mitochondria. Proc Natl Acad Sci U S A. 115(46):11796–11801.
Mereschkowsky C. 1905. Über Natur und Ursprung der Chromatophoren im Pflanzenreiche. Biol Centralbl. 25:293-604. [English translation in Martin W, Kowallik K. 1999 Annotated English translation of Mereschkowsky's 1905 paper "Über Natur und Ursprung der Chromatophoren im Pflanzenreiche." *Eur. J. Phycol.* 34, 287–295].
Mereschkowsky C. 1910. English translation in Kowallik KV, Martin WF. 2021. The origin of symbiogenesis: an annotated English translation of Mereschkowsky's 1910 paper on the theory of two plasma lineages. BioSyst. Biol. Centralbl. 30:104281. Theorie der zwei Plasmaarten als Grundlage der Symbiogenesis, einer neuen Lehre von der Entstehung der Organismen. 278–288; 289–303; 321–347; 353–367. [199].
Michod RE. 1999. Darwinian dynamics. Oxford: Oxford University Press.
Mills DB. 2020. The origin of phagocytosis in Earth history. Interface Focus. 10(4):20200019.
Mora C, Tittensor P, Adl SM, Simpson AG, Worm B. 2011. How many species are there on Earth and in the ocean? PLoS Biol. 9(8):e1001127.
Müller M, et al. 2012. Biochemistry and evolution of anaerobic energy metabolism in eukaryotes. Microbiol Mol Biol Rev. 76(2):444–495.
Nagies FSP, Brueckner J, Tria FDK, Martin WF. 2020. A spectrum of verticality across genes. PLoS Genet. 16(11):e1009200.
Niklas KJ, Cobb ED, Crawford DR. 2013. The evo-devo of multinucleate cells, tissues, and organisms, and an alternative route to multicellularity. Evol Dev. 15(6):466–474.
Nowak MA. 2006. Evolutionary dynamics: exploring the equations of life. Cambridge: Harvard University Press.
Pesole G, et al. 2012. The neglected genome. EMBO Rep. 13(6):473–474.
Pittis AA, Gabaldón T. 2016. Late acquisition of mitochondria by a host with chimaeric prokaryotic ancestry. Nature 531(7592):101–104.
Radzvilavicius AL, Blackstone NW. 2018. The evolution of individuality, revisited. Biol Rev Camb Philos Soc. 93(3):1620–1633.
Ramesh MA, Malik SB, Logsdon JM. Jr 2005. A phylogenomic inventory of meiotic genes: evidence for sex in *Giardia* and an early eukaryotic origin of meiosis. Curr Biol. 15(2):185–191.
Rand DM, Mossman JA. 2020. Mitonuclear conflict and cooperation govern the integration of genotypes, phenotypes and environments. Philos Trans R Soc B. 375(1790):20190188.
Sagan L. 1967. On the origin of mitosing cells. J Theoret Biol. 14(3):225–274.
Sánchez-Baracaldo P, Raven JA, Pisani D, Knoll AH. 2017. Early photosynthetic eukaryotes inhabited low-salinity habitats. Proc Natl Acad Sci U S A. 114(37):E7737–E7745.
Schrøder JM, et al. 2011. EB1 and EB3 promote cilia biogenesis by several centrosome-related mechanisms. J Cell Sci. 124(Pt 15):2539–2551.,
Shalev Y, Turgeman-Grott I, Tamir A, Eichler J, Gophna U. 2017. Cell surface glycosylation is required for efficient mating of *Haloferax volcanii*. Front Microbiol. 8:1253.

Shimamoto Y, Forth S, Kapoor TM. 2015. Measuring pushing and braking forces generated by ensembles of kinesin-5 crosslinking two microtubules. Dev Cell. 34(6):669–681.
Sitte P, Ziegler H, Ehrendorfer F, Bresinsky A. 1991. Strasburger, Lehrbuch der Botanik für Hochschulen. 33rd ed. Stuttgart: Gustav Fischer.
Spang A, et al. 2015. Complex archaea that bridge the gap between prokaryotes and eukaryotes. Nature. 521(7551):173–179.
Speijer D, Lukeš J, Eliáš M. 2015. Sex is a ubiquitous, ancient, and inherent attribute of eukaryotic life. Proc Natl Acad Sci U S A. 112(29):8827–8834.
Stechmann A, Cavalier-Smith T. 2002. Rooting the eukaryote tree by using a derived gene fusion. Science 297(5578):89–91.
Strasburger E. 1908. Einiges über Characeen und Amitose. Wiesner Festschrift im Auftrage Des Festkomitees Redigiert Von K. Linsbauer. Wien: C. Konegen. p. 24–47.
Szathmáry E, Demeter L. 1987. The stochastic corrector model. J Theor Biol. 128(4):463–486.
Tang Q, Pang K, Yuan X-L, Xiao S-H. 2020. A one-billion-year-old multicellular chlorophyte. Nat Ecol Evol. 4(4):543–547.
Tromer EC, van Hooff JJE, Kops GJPL, Snel B. 2019. Mosaic origin of the eukaryotic kinetochore. Proc Natl Acad Sci U S A. 116(26):12873–12882.
Tria FDK et al. 2021. Gene duplications trace mitochondria to the onset of eukaryote complexity. Genome Biol Evol. 10.1093/gbe/evab055
Vossbrinck CR, et al. 1987. Ribosomal RNA sequence suggests microsporidia are extremely ancient eukaryotes. Nature 326(6111):411–414.,
Vosseberg J, et al. 2021. Timing the origin of eukaryotic cellular complexity with ancient duplications. Nat Ecol Evol. 5(1):92–99.
Yubuki N, Leander BS. 2013. Evolution of microtubule organizing centers across the tree of eukaryotes. Plant J. 75(2):230–244.
Zaremba-Niedzwiedzka K, et al. 2017. Asgard archaea illuminate the origin of eukaryotic cellular complexity. Nature 541(7637):353–358.

Associate editor: Martin Embley Dr.

Supplementary Methods and Information

Reference tree construction

Evolutionary relationships within *Archaeplastida* are well resolved. Representatives of three major archaeplastid lineages are included – *Glaucophyta* (without subdivision), *Rhodophyta* (red algae, represented with all the major classes, and *Chloroplastida* (green algae including plants, represented with major lineages of *Streptophyta* and *Chlorophyta* (Sánchez-Baracaldo et al., 2017; Yang et al., 2016; Clerck et al., 2012). Evolution and diversity of *Amoebozoa* have not been resolved yet, nor have relationships within the supergroup; thus, we include three major groups – *Discosea*, *Tubulinea*, and *Evosea* (with evosean classes included) (Archibald et al., 2017; Adl et al., 2019; Cavalier-Smith et al., 2016), but without completely resolved relationships between them. *Excavata* monophyly is sometimes not well supported (Archibald et al., 2017; Adl et al., 2019) (see section **Tree topology and root**), but relationships within the *Discoba* and the *Metamonada* are better resolved (Kamikawa et al., 2014). Monophyly of *Hacrobia* is questionable hence unresolved (Archibald et al., 2017; Adl et al., 2019), but relationships within *Haptista* and *Cryptista* (Burki et al., 2016) appear sufficiently resolved. The supergroup of *Opisthokonta* (or *Obazoa* (Adl et al., 2019) includes *Holozoa* (animals and relatives) and *Nucletmycea* (or *Holomycota*, fungi and relatives). Relationships between the major phyla and superphyla of the animal kingdom and its closest relatives are more or less resolved (Spatafora et al., 2017). The position of *Apusomonadida* and *Breviatea* towards other opisthokonts is however still unresolved (Brown et al., 2013). SAR represents a large assemblage including three divergent groups – Stramenopiles (brown, golden-brown and yellow-green algae, diatoms, and water molds), Alveolata (ciliates, apicomplexans and dinoflagellates), and Rhizaria (including well-known radiolarians and foraminiferans) (Archibald et al., 2017; Adl et al., 2019). As opposed to more or less well-defined relations of *Alveolata* groups, taxonomy and evolution of Stramenopiles and Rhizaria are still under debate (Cavalier-Smith 2018).

Terminology of the multinucleated forms.

The terms *syncytium*, *coenocyte, multinuclear* or *polykaryon*, often used synonymously, describe organisms with numerous nuclei in shared cytoplasm. For the present paper, the terms multinucleated or multinucleate are used synonymously to describe the state of a cytoplasm with more than one nucleus, regardless of whether that state arose by fusions of cells or by nuclear division without cell division. Note that we do not score zygotes before karyogamy as

multinucleated, because otherwise all sexual lineages would be counted as having multinucleated phases.

Mathematical representation for survival probabilities of a monokaryon vs polykaryon

To compare the relative fitness of a mononucleated cell (monokaryon) and a multinucleated cell (polykaryon), we consider the difference between the probability of survival for a population of unicellular mononucleate eukaryotes *versus* that for a single syncytium. The fitness of mitochondria ($w_{mt} \in [0,1]$) and the fitness of a nucleus ($w_n \in [0,1]$) mutually affect one another ($w_{mt}w_n \in [0,1]$), such that the probability of extinction (P) of a single cell is:

$$(I)\ P = 1 - w_{mt}w_n, P \in [0,1]$$

For a population of monokaryons, the probability of population extinction is the product of the probabilities of extinction for each cell composing the population, so that (I) becomes:

$$(II)\ P_{population} = \prod P = \prod (1 - w_{mt}w_n)\,, P_{population} \in [0,1]$$

For a syncytium, the probability of extinction is not comparable to the survival probability for a monokaryon, but to a population of monokaryons:

$$(III) P_{syncytium} = 1 - w_{mt(syncytium)}w_{n(syncytium)}, P_{syncytium} \in [0,1]$$

However, $w_{mt(syncytium)}$ and $w_{n(syncytium)}$ are not expressible in terms of $w_n$and w_{mt} in the monokaryon population. Why is the fitness more comparable between the monokaryon population and a syncytium than between a single monokaryon cell and a syncytial cell? This is schematically shown in **Fig. 4**. A syncytium behaves as more like a population of nuclei and mitochondria than as an individual cell. Thus, if $(1 - w_{mt(syncytium)}w_{n(syncytium)}) \geq P_{population}$, then the syncytium has a higher chance of survival than a population of monokaryons.

Supplementary References

Adl SM et al., 2019. Revisions to the classification, nomenclature, and diversity of eukaryotes. *J Eukaryot Microbiol*. 66:4-119.

Archibald JM, Simpson AGB, Slamovits CH, editors. 2017. Handbook of the Protists. *Springer Nature*.

Brown MW et al., 2013. Phylogenomics demonstrates that breviate flagellates are related to opisthokonts and apusomonads. *Proc R Soc B* 280:20131755

Burki F et al., 2016. Untangling the early diversification of eukaryotes: A phylogenomic study of the evolutionary origins of Centrohelida, Haptophyta and Cryptista. *Proc R Soc B* 283:20152802.

Cavalier-Smith T. 2018. Kingdom Chromista and its eight phyla: a new synthesis emphasising periplastid protein targeting, cytoskeletal and periplastid evolution, and ancient divergences. *Protoplasma*. 255:297-357.

Cavalier-Smith T, Chao EE, Lewis R. 2016. 187-gene phylogeny of protozoan phylum Amoebozoa reveals a new class (Cutosea) of deep-branching, ultrastructurally unique, enveloped marine Lobosa and clarifies amoeba evolution. *Mol Phylogenet Evol*. 99:275– 296.

Clerck OD, Bogaert KA, Leliaert F. 2012. Diversity and Evolution of Algae: Primary Endosymbiosis. *Adv Bot Res*. 64:55–86.

Kamikawa R et al. 2014. Gene content evolution in discobid mitochondria deduced from the phylogenetic position and complete mitochondrial genome of *Tsukubamonas globosa*. *Genome Biol Evol*. 6:306–315.

Sánchez-Baracaldo P, Raven JA, Pisani D, Knoll AH. 2017. Early photosynthetic eukaryotes inhabited low-salinity habitats. *Proc Natl Acad Sci USA*. 114:E7737-E7745.

Spatafora et al. 2017. The Fungal Tree of Life: From Molecular Systematics to Genome-Scale Phylogenies. In: Heitman et al. (Eds.), The Fungal Kingdom. Washington: ASM Press. Pp. 3-34.

Yang EC et al., 2016. Divergence time estimates and the evolution of major lineages in the florideophyte red algae. *Sci Rep*. 6:21361.

Supplementary Table 1: Summary of eukayarotic groups analysed.

TAXON	RANK	G	SG	HG	NUMBER OF KNOWN SPECIES	NUMBER OF SEQUENCES SPECIES	LIST OF SEQUENCED SPECIES
DINOFLAGELLATA	phylum				2294	1	*Symbiodinium minutum*
PERKINSOZOA	phylum				10	N/A	*N/A*
APICOMPLEXA	phylum	Alveolata			4500	10	*Babesia bigemina, B. bovis, Cryptosporidium parvum, Neospora caninum, Plasmodium chabaudi, P. falciparum, P. vivax, Toxoplasma gondii, Theileria equi, T. parva*
CHROMERIDA	phylum				2	2	*Chromera velia, Vitrella brassicaformis*
CILIOPHORA	phylum				3500	2	*Oxytricha trifallax, Tetrahymena thermophila*
XANTHOPHYCEAE	class	Stramenopiles	SAR		688	N/A	*N/A*
PHAEOPHYCEAE	class				2045	1	*Ectocarpus siliculosus*
RAPHIDOPHYCEAE	class				40	N/A	*N/A*
CHRYSOPHYTA	class				736	N/A	*N/A*
EUSTIGMATOPHYCEAE	class				105	1	*Nannochloropsis gaditana*
SYNCHROMOPHYCEAE	class				100	N/A	*N/A*
PELAGOPHYCEAE	class				25	1	*Aureococcus anophagefferens*
BACILLARIOPHYCEAE	class				15561	2	*Phaeodactylum tricornutum, Thalassiosira pseudonana*
ACTINOPHRYIDA	class				16	N/A	*N/A*
OOMYCOTA	order				520	2	*Saprolegnia parasitica, Phytophthora sojae*
HYPHOCHYTRIOMYCOTA	class				11	N/A	*N/A*
BLASTOCYSTIDA	class				10	1	*Blastocystis hominis*
OPALINATA	class				10	N/A	*N/A*
LABYRINTHULOMYCOTA	class				85	N/A	*N/A*
FORAMINIFERA	subphylum	Rhizaria			50000	1	*Reticulomyxa filosa*
POLYCYSTINEA	class				800	N/A	*N/A*
ACANTHAREA	class				260	N/A	*N/A*
PARAMYXIDA	order				13	N/A	*N/A*
HAPLOSPORIDIA	order				55	N/A	*N/A*
PHYTOMYXEA	class				35	1	*Plasmodiophora brassicae*
VAMPYRELLIDAE	family				19	N/A	*N/A*
FILOSA	class				1313	1	*Bigelowiella natans*
COCCOLITHALES	order	Haptista	Hacrobia	DIAPHORETICKES	148	N/A	*N/A*
ISOCHRYSIDALES	order				44	1	*Emiliania huxleyi*
PRYMNESIALES	order				83	N/A	*N/A*
PHAEOCYSTALES	order				10	N/A	*N/A*
PAVLOVALES	order				13	N/A	*N/A*
CRYPTOPHYCEAE	class	Cryptista			220	1	*Guillardia theta*
GONIOMONADIDA	class				5	N/A	*N/A*
KATABLEPHARIDEA	class				16	N/A	*N/A*
FLORIDEOPHYCEAE	class	Rhodophyta			6724	1	*Chondrus crispus*
BANGIOPHYCEAE	class				194	N/A	*N/A*
RHODELLOPHYCEAE	class				6	N/A	*N/A*
COMPSOPOGONOPHYCEAE	class				75	N/A	*N/A*
STYLONEMATOPHYCEAE	class				39	N/A	*N/A*
PORPHYRIDIOPHYCEAE	class				12	1	*Porphyridium purpureum*
CYANIDIOPHYCEAE	class				7	2	*Cyanidioschyzon merolae, Galdieria sulphuraria*
EUPHYLLOPHYTINA	superphylum	Chloroplastida	Archaeplastida		310000	7	*Ananas comosus, Arabidopsis thaliana, Musa acuminata, Oryza sativa, Phaseolus vulgaris, Sorghum bicolor, Zea mays*
LYCOPODIOPHYTA	phylum				1300	1	*Selaginella moellendorffii*
ANTHOCERATOPHYTA	phylum				300	N/A	*N/A*
BRYOPHYTA	phylum				12000	1	*Physcomitrella patens*
MARCHANTIOPHYTA	phylum				9000	N/A	*N/A*
ZYGNEMATOPHYTA	phylum				3900	N/A	*N/A*
COLEOCHAETOPHYCEAE	class				36	N/A	*N/A*
CHAROPHYCEAE	class				700	N/A	*N/A*
KLEBSORMIDIOPHYCEAE	class				23	N/A	*N/A*
MESOSTIGMA	*genus*				2	N/A	*N/A*
CHLOROKYBOPHYCEAE	class				1	N/A	*N/A*
ULVOPHYCEAE	class				1900	N/A	*N/A*
CHLOROPHYCEAE	class				3600	2	*Chlamydomonas reinhardtii, Volvox carteri*
TREBOUXIOPHYCEAE	class				885	2	*Coccomyxa subellipsoidea, Chlorella variabilis*
MAMIELLALES	order				22	4	*Bathycoccus prasinos, Micromonas commoda, Ostreococcus lucimarinus, O. tauri*
MONOMASTIGALES	order				6	N/A	*N/A*
PALMOPHYLLALES	order				6	N/A	*N/A*
PRASINOCOCCALES	order				3	N/A	*N/A*
GLAUCOPHYTA	phylum	- II -			23	1	*Cyanophora paradoxa*
JAKOBIDA	order	Discoba	'Excavata'	EXCAVATA	20	N/A	*N/A*
HETEROLOBOSEA	class				150	1	*Naegleria gruberi*
EUGLENIDA	phylum				1500	N/A	*N/A*
DIPLONEMIDA	order				9	N/A	*N/A*
KINETOPLASTEA	class				90	5	*Leishmania donovani, L. infantum, L. panamensis, Trypanosoma b. brucei, T. brucei gambiense*
DIPLOMONADIDA	order	Metamonada			65	1	*Giardia lamblia*
RETORTAMONAS	*genus*				30	N/A	*N/A*
DYSNECTES	*genus*				1	N/A	*N/A*

KIPFERLIA	*genus*				1	N/A	*N/A*
CHILOMASTIX	*genus*				30	N/A	*N/A*
CAVIOMONADIDAE	family				8	N/A	*N/A*
PARABASALIA	class				450	1	*Trichomonas vaginalis*
PREAXOSTYLA	class				150	N/A	*N/A*
DEUTEROSTOMIA	superphylum				110000	30	*Anolis carolinensis, Bos indicus, Callithrix jacchus, Canis lupus familiaris, Chrysemys picta bellii, Ciona intestinalis, Coturnix japonica, Cynoglossus semilaevis, Cyprinus carpio, Equus caballus, Esox lucius, Felis catus, Gallus gallus, Gorilla g. gorilla, Homo sapiens, Ictalurus punctatus, Lepisosteus oculatus, Macaca mulatta, Monodelphis domestica, Mus musculus, Nothobranchius furzeri, Oreochromis niloticus, Oryctolagus cuniculus, Oryzias latipes, Ovis aries, Pongo abelii, Pan paniscus, Salmo salar, Taeniopygia guttata, Xenopus laevis*
PROTEROSTOMIA	superphylum		**Holozoa**		1500000	16	*Aedes aegypti, Anopheles gambiae, Apis mellifera, Bombus terrestris, Caenorhabditis briggsae, C. elegans, Crassostrea virginica, Drosophila busckii, D. melanogaster, D. miranda, D. p. pseudoobscura, D. simulans, D. yakuba, Nasonia vitripennis, Schistosoma mansoni, Tribolium castaneum*
CNIDARIA	phylum				11000	N/A	*N/A*
CTENOPHORA	phylum				150	N/A	*N/A*
PLACOZOA	phylum				1	1	*Trichoplax adhaerens*
HOMOSCLEROMORPHA	class				120	N/A	*N/A*
CALCAREA	class				400	N/A	*N/A*
HEXACTINELLIDA	class				1300	N/A	*N/A*
DEMOSPONGIAE	class				8800	1	*Amphimedon queenslandica*
CHOANOFLAGELLATA	class				150	1	*Monosiga brevicollis*
FILASTEREA	class		**Obazoa**		5	1	*Capsaspora owczarzaki*
MESOMYCETOZOEA	class				60	N/A	*N/A*
CORALLOCHYTRIUM	*genus*				1	N/A	*N/A*
BREVIATEA	class	**- I I -**			4	N/A	*N/A*
APUSOMONADIDA	order	**- I I -**		**AMORPHEA**	22	N/A	*N/A*
CRISTIDISCOIDEA	class				25	N/A	*N/A*
MICROSPORIDIA	phylum				1500	4	*Encephalitozoon cuniculi, E. hellem, E. intestinalis, E. romaleae*
NEOCALLIMASTIGOMYCOTA	phylum				34	2	*Orpinomyces sp., Piromyces sp.*
CHYTRIDIOMYCOTA	phylum				1090	4	*Batrachochytrium dendrobatidis, Gonapodya prolifera, Homolaphlyctis polyrhiza, Spizellomyces punctatus*
BLASTOCLADIOMYCOTA	phylum				120	N/A	*N/A*
ZOOPAGOMYCOTINA	phylum				250	7	*Basidiobolus meristosporus, Coemansia reversa, Conidiobolus coronatus, Linderina pennispora, Martensiomyces pterosporus, Piptocephalis cylindrospora, Ramicandelaber brevisporus*
MUCOROMYCOTA	phylum		**Nucletmycea**		325	7	*Backusella circina, Hesseltinella vesiculosa, Lichtheimia corymbifera, Mortierella verticillata, Rhizopus delemar, Saksenaea vasiformis, Umbelopsis ramanniana*
BASIDIOMYCOTA	phylum				32000	5	*Cryptococcus gattii, C. neoformans var. grubii, C. n. var. neoformans B-3501A, Cryptococcus n. var. neoformans JEC21, Ustilago maydis*
ASCOMYCOTA	phylum				65000	9	*Aspergillus fumigatus, Fusarium graminearum, Magnaporthe oryzae, Neurospora crassa, Pochonia chlamydosporia, Saccharomyces cerevisiae, Schizosaccharomyces pombe, Thermothelomyces thermophila, Zymoseptoria tritici*
MYXOGASTRIA	class				1000	N/A	*N/A*
PROTOSPORANGIIDA	order				28	N/A	*N/A*
DICTYOSTELIA	class		**Evosea**	**Amoebozoa**	55	2	*Dictyostelium discoideum, Polysphondylium pallidum*
VARIOSEA	class				44	N/A	*N/A*
CUTOSEA	class				4	N/A	*N/A*
ARCHAMOEBEA	class				36	1	*Entamoeba histolytica*
TUBULINEA	class	**- I I -**			480	N/A	*N/A*
DISCOSEA	class	**- I I -**			180	1	*Acanthamoeba castellanii*

Supplementary Table 2: Meta-data with all eukaryotic taxa and accompanying traits.
Shown are supergroup, higher group, included taxa, corresponding taxonomic rank (phylum, class, order, family or genus), as well as the higher ranks to which each taxon belongs (HG – highest group, SG – supergroup, G- group). Annotated is also number of known species per each group, as well as species for which genomes are sequenced.

TAXON	MULTINUCLEATED EXAMPLES	SEX	POLYPLOID (2N OR MORE)	TYPE OF MITOSIS	PLASTID	MITOCHONDRIA
DINOFLAGELLATA	*Syndiniophyceae*	present	Some species may have diploid generation	Closed extranuclear pleuromitosis	Secondary, red alga-derived, or tertiary, diatom-derived, haptophyte-derived, or green-alga derived	Mitochondria with ampulliform (type of tubular) cristae
PERKINSOZOA	Young sporocytes of most representatives	absent	Not known	Closed extranuclear pleuromitosis	Absent	Mitochondria with tubular cristae
APICOMPLEXA	Schizont phase in almost all the species	present	Temporary polyploid schizont nucleus	Open or semi-open pleuromitosis or orthomitosis	Photosynthesis lost (apicoplast), originally red alga, probably also surrounded by four membranes	Mitosomes (*Cryptosporidium*) or mitochondria in *Plasmodium*
CHROMERIDA	no	not known	Not known	Closed extranuclear mitosis	Secondary, from red alga; single peripherally located and prolonged chloroplast	Mitochondria with tubular cristae
CILIOPHORA	*Pseudokeronopsis*, Litostomatea: Spathidiidae, *Dileptus*, Intromacronucleata	present	Some species have huge polyploidy number	Closed orthomitosis	Recent examples of tertiary (*Myrionecta rubra, Tontonia appendiculariformis*, *Paramecium bursaria*)	Mitochondria or hydrogenosomes
XANTHOPHYCEAE	*Vaucheria, Tribonematales*	present	There are species with 2n dominant	Closed orthomitosis	Secondary, from red alga; rarely one or two, usually many	Mitochondria with tubular cristae
PHAEOPHYCEAE	*Haplospora* has multinucleated not motile spores; multinucleated cells reported, but also all the cells rich in plasmodesmata connections	present	Some with diploidy dominant, polyploids >2n	Closed orthomitosis	Secondary, from red alga; one to many; many discoid and without pyrenoids	Mitochondria with tubular cristae
RAPHIDOPHYCEAE	No	present	Diploidy dominant	Semi-open extranuclear orthomitosis	Secondary, from red alga; multiple plastids	Mitochondria with tubular cristae
CHRYSOPHYTA	*Poterioochromonas* on plates, *Botrydiopsis pyrenoidosa*	present	Not known	Open orthomitosis	Secondary, from red alga; usually one or two in a cell	Mitochondria with tubular cristae
EUSTIGMATOPHYCEAE	No	not known	A lot of species with 2n dominant	Not known	Secondary, from red alga	Mitochondria with tubular cristae
SYNCHROMOPHYCEAE	*Chlamydomyxa*, aggregates with multiple nuclei	present	A lot of species with 2n dominant	(probably closed)	Secondary, from red alga; chloroplast in aggregates – six to eight of them	Mitochondria with tubular cristae
PELAGOPHYCEAE	No	not known	A lot of species with 2n dominant	Not known	Secondary, from red alga	Mitochondria with tubular cristae
BACILLARIOPHYCEAE	Structures analogous to multinucleated structures occur in epilithon species	present	Diploidy dominant, some >2n, *Phaeodactylum tricornutum*	Closed or in rare species open orthomitosis	Secondary, from red alga; one to many, bounded by four membranes, discoid, lobed, plate- or ribbon-like	Mitochondria with tubular cristae
ACTINOPHRYIDA	Several peripheral nuclei	present	A lot of species with 2n dominant	Semi-open orthomitosis	Absent	Mitochondria with tubular cristae
OOMYCOTA	Aseptate thallus	present	Known in *Phytophthora*	Closed intranuclear pleuromitosis	Absent	Mitochondria with tubular cristae
HYPHOCHYTRIOMYCOTA	Binucleate to multinucleate cells	not known	A lot of species with 2n dominant	Closed intranuclear pleuromitosis	Absent	Mitochondria with tubular cristae
BLASTOCYSTIDA	Multinucleated cells	not known	Not known	Closed intranuclear pleuromitosis	Absent	Mitochondria with tubular cristae
OPALINATA	*Cepedea, Opalina, Protoopalina*; two to many monomorphic nuclei; Opalinidae have two to many nuclei	present	A lot of species with 2n dominant	Closed intranuclear mitosis	Absent	Mitochondria with tubular cristae
LABYRINTHULOMYCOTA	Numerous colonial examples, mononucleate unicellulars; colonies analogous to coenocyte, but every cell separated in e.g. *Aplanochytrium*, genera of Labyrinthulidae connected with bothrosome to ectoplasmic network; vegetative cells with more than one nucleus	not known	A lot of species with 2n dominant	Open orthomitosis	A lot of *Amphitremidaceae* have numerous *Trebouxiophycean* endosymbionts and exhibit mixotrophy	Mitochondria with tubular cristae
FORAMINIFERA	Monothalamea (including Xenophyophorea), and many others, especially deeps sea taxa	present	Some species polyploid (>2n)	Closed intranuclear pleuromitosis	Absent	Mitochondria aerobic, with tubular cristae
POLYCYSTINEA	*Collozoum*	absent	Not known	(probably closed intranuclear mitosis)	Absent	Mitochondria aerobic, with tubular cristae
ACANTHAREA	Adults usually multinucleated	present	Not known	(probably closed intranuclear mitosis)	Absent	Mitochondria aerobic, with flattened cristae
PARAMYXIDA	Multicellular/ multinucleated spore, first cell has its nucleus and one spore in the cytosol, second has its nucleus and 6 spores inside of its cytosol, of which each has nucleus = multinucleated spore?	not known	Maybe diploids, since lacking sex (?)	(probably closed intranuclear mitosis)	Absent	Mitochondria with tubular cristae
HAPLOSPORIDIA	Multinucleate plasmodium	not known	Polyploidy is multiplication of 3	Closed intranuclear pleuromitosis	Absent	Mitochondria with tubular cristae

PHYTOMYXEA	Primary and secondary plasmodium multinucleated	present	Diploid phase?	Closed intranuclear pleuromitosis or orthomitosis	Absent	Mitochondria with tubular cristae
VAMPYRELLIDAE	Plasmodia originated from cell fusion	not known	Not known	Closed orthomitosis	Absent	Mitochondria with tubular cristae
FILOSA	Multinucleate syncytium of *Chlorarachniophyceae*, and many more, especially deep-sea taxa	present	Nucleomorph polyploid, nucleus haploid	Open orthomitosis in *Heliomorpha*, in others closed intranuclear orthomitosis	Secondary (for example *Chlorarachniophyceae*)	Mitochondria with tubular cristae
COCCOLITHALES	No	present	Diploid phase of lifecycle	Open orthomitosis	Secondary; usually one or two	Mitochondria with tubular cristae, single highly branched mitochondrion
ISOCHRYSIDALES	No	present	Diploid phase of lifecycle	Open orthomitosis	Secondary, from red alga; one to four (four membranes known only in *Dicrateria*)	Mitochondria with tubular cristae, single highly branched mitochondrion
PRYMNESIALES	No	present	Diploid phase of lifecycle	Open orthomitosis	Secondary; two per cell	Mitochondria with tubular cristae, single highly branched mitochondrion
PHAEOCYSTALES	Nonmotile cells colonial and embedded in gelatinous matrix, functionally syncytium	present	Probably one stage in the life cycle diploid, while other haploid, or only zygote being diploid	Open orthomitosis	Secondary; one to four chloroplast per cell	Mitochondria with tubular cristae, single highly branched mitochondrion
PAVLOVALES	No	not known	Not known	Semi-open orthomitosis	Secondary; usually single; but sometimes two	Mitochondria with tubular cristae, single highly branched mitochondrion
CRYPTOPHYCEAE	No	not known	Nucleomorph polyploid, nucleus haploid	Open orthomitosis	Secondary	Mitochondrial cristae flat tubules
GONIOMONADIDA	No	not known	Diploidy maybe dominant?	Not known	Absent	Mitochondrial cristae flat tubules
KATABLEPHARIDEA	No	not known	Not known	Not known	Absent	Mitochondria with tubular cristae
FLORIDEOPHYCEAE	Most Florideophyceae have multinucleated thalli	present	Diploid phases of lifecycle	Closed mitosis	Primary; numerous in multinucleate cells	Mitochondria with flat cristae
BANGIOPHYCEAE	No	present	Diploid phases of lifecycle	Closed mitosis	Primary; single stellate in the center of the cell, or sometimes many peripheral	Mitochondria with flat cristae
RHODELLOPHYCEAE	No	not known	Not known	Closed mitosis	Primary; single highly lobed	Mitochondria with flat cristae
COMPSOPOGONOPHYCEAE	Multinucleated thallus forming spores	present	Diploid phases of lifecycle	Closed mitosis	Primary; numerous in each cell	Mitochondria with flat cristae
STYLONEMATOPHYCEAE	Loose colonies of cells in shared mucus, functionally syncytium	not known	Possibly diploids present	Closed mitosis	Primary; single stellate	Mitochondria with flat cristae
PORPHYRIDIOPHYCEAE	No	present	Diploidy known	Closed or semi-open mitosis	Primary; single, branched or stellate	Mitochondria with flat cristae
CYANIDIOPHYCEAE	species with multiple nuclei known	absent	Not known	Closed mitosis	Primary	Mitochondria with flat cristae
EUPHYLLOPHYTINA	Endosperm, tapetal cells, latex cells, storage cells	present	Dominant 2n; polyploids known	Open orthomitosis	Primary	Mitochondria with flat cristae
LYCOPODIOPHYTA	Some epidermal and cortical cells multinucleate	present	Dominant 2n, gametophyte n	Open orthomitosis	Primary	Mitochondria with flat cristae
ANTHOCERATOPHYTA	Multinucleate cells present	present	Dominant n (gametophyte), sporophyte 2n	Open orthomitosis	Primary	Mitochondria with flat cristae
BRYOPHYTA	Multinucleate cells present	present	Dominant n (gametophyte), sporophyte 2n	Open orthomitosis	Primary	Mitochondria with flat cristae
MARCHANTIOPHYTA	Multinucleate cells present	present	Dominant n (gametophyte), sporophyte 2n	Open orthomitosis	Primary	Mitochondria with flat cristae
ZYGNEMATOPHYTA	In some species, cells not clearly delimited, pit connections wide	present	Numerous species polyploid	Open orthomitosis	Primary; one to many, coming in various shapes	Mitochondria with flat cristae
COLEOCHAETOPHYCEAE	No	present	Not known (only zygote diploid)	Open orthomitosis	Primary	Mitochondria with flat cristae
CHAROPHYCEAE	Multinucleate internodal cells; structures analogous to phloem	present	Some polyploid, but thallus haploid, polyploidy adaptive in self-fertilizing species	Open orthomitosis	Primary; multinucleated cells with numerous oval plastids	Mitochondria with flat cristae
KLEBSORMIDIOPHYCEAE	No	not known	Not known	Open orthomitosis	Primary; one or two chloroplasts	Mitochondria with flat cristae
MESOSTIGMA	No	not known	Not known	Open orthomitosis	Primary; single chloroplast in a cell	Mitochondria with flat cristae
CHLOROKYBOPHYCEAE	No	absent	Not known	Open orthomitosis	Primary, single chloroplast in a cell	Mitochondria with flat cristae
ULVOPHYCEAE	Multinucleated siphonous thallus	present	Diploid phase exists	Closed mitosis	Primary, many in shared coenocyte	Mitochondria with flat cristae
CHLOROPHYCEAE	*Scenedesmus*, *Hydrodictyon*, and *Pediastrum*	present	Either only zygote haploid or there is 2n phase	Closed or semi-open orthomitosis	Primary; one or more	Mitochondria with flat cristae
TREBOUXIOPHYCEAE	No	not known	Not known	Closed or semi-open mitosis	Primary	Mitochondria with flat cristae
MAMIELLALES	No	not known	Not known	Not known	Primary; one or two per cell	Mitochondria with flat cristae
MONOMASTIGALES	No	not known	Not known	Not known	Primary or absent	Mitochondria with flat cristae
PALMOPHYLLALES	Aggregation of unicellulars in common loose matrix, functionally syncytium	not known	Not known	Not known	Primary; single cup shaped	Mitochondria with flat cristae
PRASINOCOCCALES	Some species forming loose colonies, in some species maybe multinucleated 'tissue'	absent	Not known	Closed intranuclear pleuromitosis	Primary; single cup shaped	Mitochondria with flat cristae

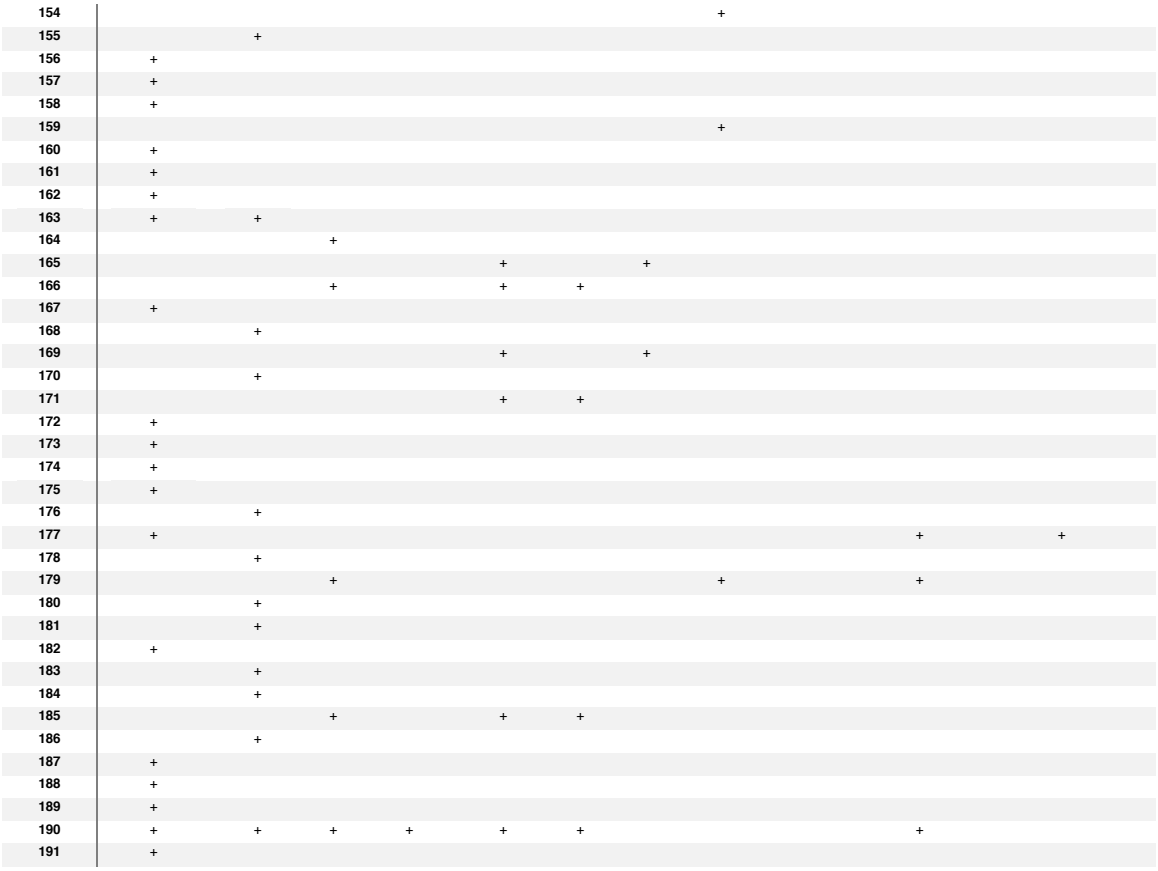
GLAUCOPHYTA	*Cyanophora paradoxa* on plate	absent	Not known	Open mitosis	Primary; mucoroplast with peptidoglycan wall, usually two or more	Mitochondria with flat cristae
JAKOBIDA	No	not known	Not known	Open mitosis	Absent	Aerobic with tubular mitochondrial cristae or anaerobic with acristate mitochondria
HETEROLOBOSEA	Gruberellidae, Acrasidae	not known	Polyploidy known	Closed mitosis	Absent	Mitochondrial cristae flattened, often discoidal, sometimes acristate
EUGLENIDA	No	present	Not known	Closed orthomitosis	Secondary, *Pyramomonadales* – related	Mitochondria with discoid (paddle-shaped cristae)
DIPLONEMIDA	No	not known	Not known	Closed orthomitosis	Absent	Mitochondria with giant flattened discoidal cristae
KINETOPLASTEA	No	present	Diploidy dominant, aneuploidy observed	Closed intranuclear pleuromitosis or orthomitosis	Absent	Mitochondria with discoid or tubular cristae
DIPLOMONADIDA	Most members binucleated	present	Tetraploidy dominant	Semi-open mitosis	Absent	Hydrogenosome (*Spironucleus*) or mitosome (*Giardia*)
RETORTAMONAS	No	absent	Not known	Closed mitosis	Absent	not known
DYSNECTES	No	absent	Not known	Not known	Absent	Hydrogenosome-like
KIPFERLIA	No	absent	Diploidy known	Not known	Absent	Hydrogenosome-like
CHILOMASTIX	No	absent	Not known	Closed mitosis	Absent	Hydrogenosome-like
CAVIOMONADIDAE	No	absent	Not known	Not known	Absent	Hydrogenosome-like
PARABASALIA	Some Tritrichomonadida binucleated (in permanent telophase), Cristamonadida; multinucleated forms with observed division into unequal daughter cells, with unequal number of nuclei	present	Not known	Closed extranuclear pleuromitosis	Absent	Hydrogenosomes
PREAXOSTYLA	*Microrhopalodina*, *Barroella*	present	Diploidy or tetraploidy dominant	Open or closed intranuclear pleuromitosis	Absent	Hydrogenosomes or absent
DEUTEROSTOMIA	Mammal placenta is syncytial, skeletal muscles of mammals and other vertebrates are syncytial, as well as osteoclasts and chondroclasts	present	Diploidy dominant, more than 2n rare	Open mitosis	Absent	Mitochondria with flat cristae
PROTEROSTOMIA	Nematoda have examples of multinucleated cells and/or tissues, *Drosophila* larva, and probably most of holometabolous insects have syncytial larva, syncytial neodermis in Platyhelminthes	present	Diploidy dominant, haploid and males in some species	Open or semi-open mitosis	Absent	Mitochondria with flat cristae
CNIDARIA	Myxozoa binucleated spores or coenocytes	present	Diploidy dominant	Closed intranuclear pleuromitosis in Myxozoa	Absent	Mitochondria with flat cristae
CTENOPHORA	Smooth muscle fibers, Myxozoa have multinucleated plasmodia	present	Diploidy dominant	Open mitosis	Absent	Mitochondria with flat cristae
PLACOZOA	*Trichoplax* contractile muscle cells	not known	Diploidy dominant	Not known	Absent	Mitochondria with flat cristae
HOMOSCLEROMORPHA	No	present	Diploidy dominant	Open or semi-open mitosis	Absent	Mitochondria with flat cristae
CALCAREA	*Scypha ciliata* blastomeres have multiple nuclei	present	Diploidy dominant	Open or semi-open mitosis	Absent	Mitochondria with flat cristae
HEXACTINELLIDA	Body is a single continuous syncytium	present	Diploidy dominant	Not known	Absent	Mitochondria with flat cristae
DEMOSPONGIAE	*Ephydatia fluviatilis* gemules, *Halichondria panicea* multinucleate spermatozoa	present	Diploidy dominant	Open mitosis	Absent	Mitochondria with flat cristae
CHOANOFLAGELLATA	No	present	Alternation of n and 2n	Open or semi-open mitosis	Absent	Mitochondria with flat, not discoid cristae
FILASTEREA	Aggregative phase	not known	N/A	Not known	Absent	Mitochondria with flat cristae
MESOMYCETOZOEA	*Creolimax*, multinucleated filaments	absent	N/A	Closed mitosis	Absent	Mitochondria with flat cristae, but some may have tubular cristae
CORALLOCHYTRIUM	Multinucleated schizont prior to the sporulation of amoebae	not known	N/A	Not known	Absent	Not known
BREVIATEA	*Breviata*, multinucleated amoebae	absent	N/A	Not known	Absent	Hydrogenosome-like
APUSOMONADIDA	*Multimonas* syncytia	absent	N/A	Not known	Absent	Mitochondria with tubular cristae
CRISTIDISCOIDEA	Polynucleate genera known	not known	Not known	(probably closed mitosis)	Absent	Mitochondrial cristae discoid or flat
MICROSPORIDIA	Multinucleate sporogonial plasmodium, meiosis followed by several nuclear divisions resulting in a cell with 16-32 nuclei	present	Not known	Closed intranuclear pleuromitosis	Absent	Mitosomes

NEOCALLIMASTIGOMYCOTA	Coenocytic sporangium	not known	Not known	Closed mitosis	Absent	Hydrogenosomes
CHYTRIDIOMYCOTA	Coenocytic hyphae	present	Not known	Closed or semi-open mitosis	Absent	Mitochondria with flat cristae
BLASTOCLADIOMYCOTA	Coenocytic hyphae	present	Evolutionary young polyploids known	Closed intranuclear pleuromitosis	Absent	Mitochondria with flat cristae
ZOOPAGOMYCOTA	Coenocytic hyphae	present	Not known	Closed intranuclear pleuromitosis	Absent	Mitochondria with flat cristae
MUCOROMYCOTA	Coenocytic hyphae	present	Some *Glomeromycota* have polyploid nuclei	Closed intranuclear pleuromitosis	Absent	Mitochondria with flat cristae
BASIDIOMYCOTA	Coenocytic hyphae	present	*Cyathus*	Closed intranuclear pleuromitosis	Absent	Mitochondria with flat cristae
ASCOMYCOTA	Coenocytic hyphae	present	Some species diploid, *Schizosaccharomyces pombe*, some yeasts young polyploid, like *Rhizopus*	Closed intranuclear pleuromitosis	Absent	Mitochondria with flat cristae
MYXOGASTRIA	Multinucleated plasmodium	present	Not known	Closed intranuclear orthomitosis (open in monokaryons) or pleuromitosis	Absent	Mitochondria with tubular cristae (ramicristate)
PROTOSPORANGIIDA	Protosporangiidae, *Ceratyomyxa*, multinucleated plasmodium	absent	Not known	Closed intranuclear mitosis	Absent	Mitochondria with tubular cristae (ramicristate)
DICTYOSTELIA	Agreggative syncytium, grex phase	present	Not known	Closed intranuclear mitosis	Absent	Mitochondria with tubular cristae (ramicristate)
VARIOSEA	Schizoplasmodidae, Cavosteliida, *Darvyshirella*, *Dictyamoeba*, *Arboramoeba*, *Heliamoeba*	present	Not known	Closed intranuclear mitosis	Absent	Mitochondria with tubular cristae (ramicristate)
CUTOSEA	No	absent	Not known	Closed intranuclear mitosis	Absent	Mitochondria with tubular cristae (ramicristate)
ARCHAMOEBEA	*Tricholimax*, *Pelomyxa*	absent	Polyploids known	Closed intranuclear mitosis	Absent	Mitosomes, hydrogenosomes or lacking mitochondria
TUBULINEA	*Chaos, Parachaos*	present	Polyploids known	Semi-open or closed orthomitosis	Absent	Mitochondria with tubular cristae (ramicristate)
DISCOSEA	*Acanthamoeba*, *Cochliopodium*	present	Polyploids known	Open and probably in some species closed orthomitosis	Absent	Mitochondria with tubular cristae; some with flat hydrogen producing mitochondria

Supplementary Table 3: References used for selection of taxa, reconstruction of the phylogenetic relationship, and annotation of traits.

REF	PHYLOGENY	SYNCYTIAL	SEX	ASEXUAL	POLYPLOID	2N	>2N	MITOSIS	MITOCHONDRIA	PLASTID	SECONDARY PLASTID
1		+			+				+		
2	+	+	+	+	+			+	+	+	+
3				+	+	+	+				
4		+									
5	+										
6	+	+	+	+	+	+	+	+	+	+	+
7	+										
8	+										
9		+	+	+							
10		+									
11		+									
12		+									
13		+									
14	+										
15	+										
16		+									
17		+	+	+	+	+					
18					+						
19	+										
20	+										
21					+		+				
22			+		+	+					
23			+	+							
24		+									
25			+								
26	+										
27								+			
28		+	+		+	(+)					
29	+										
30	+										
31		+									
32		+	+								
33								+	+		
34	+										
35	+										
36	+										
37	+										
38	+										
39									+		
40			+								
41			+							+	
42		+	+								
43	+										
44		+									
45								+			
46	+										
47								+			
48	+										
49			+								
50		+									
51		+									
52										+	+
53	+									+	+
54			+								
55			+								
56				+	+	+					
57	+										
58		+									
59	+										
60		+									
61		+	+					+	+		
62			+		+	+					
63					+		+				
64					+		+				
65		+									
66		+									
67		+			+		+			+	
68		+	+		+	+	+	+	+		
69		+									
70		+									
71	+										
72		+									
73	+										

74		+									
75			+								
76		+									
77					+		+ (nucleo morph)				
78	+										
79			+	+							
80	+										
81			+	+						+	+
82		+									
83	+										
84	+										
85			+								
86		+									
87	+										
88	+										
89	+										
90	+										
91		+	+		+	+				+	+
92		+	+					+		+	+
93		+						+			
94	+										
95			+							+	+
96			+					+	+		
97	+										
98								+	+	+	+
99	+									+	
100	+									+	
101	+										
102		+	+	+							
103		+	(+)	+	+		+				
104								+			
105		+	+	+	+	+			+		
106		+									
107	+										
108		+									
109	+										
110		+									
111	+										
112			+								
113	+										
114									+		
115								+			
116		+									
117			+		+	+					
118		+									
119	+										
120	+										
121								+		+	+
122					+						
123									+		
124		+								+	+
125			+								
126	+										
127	+										
128	+										
129	+										
130		+	+								
131		+	+								
132	+										
133								+			
134								+		+	
135	+										
136								+			
137		+									
138					+		+				
139					+			+			
140		+									
141		+						+	+		
142	+										
143		+									
144									+		
145		+									
146		+									
147		+									
148		+									
149		+							+		
150								+			
151	+										
152	+										
153	+										



References

1. R. D. Adam, Diplomonadida. In *Handbook of the Protists*, J. M. Archibald, A. G. B. Simpson, & C. H. Slamovits, Eds. (Springer Nature, 2017) pp.1219-1246.

2. S. M. Adl, *et al.*, Revisions to the Classification, Nomenclature, and Diversity of Eukaryotes. *The Journal of Eukaryotic Microbiology*, 66, 4-119 (2019).

3. W. Albertin, & P. Marullo, Polyploidy in fungi: evolution after whole-genome duplication. *Proceedings of the Royal Society B: Biological Sciences*, 279, 2497-2509 (2012).

4. O. R. Anderson, Ultrastructure of a colonial radiolarian *Collozoum inerme* and a cytochemical determination of the role of its zooxanthellae. *Tissue and Cell*, 8, 195-208 (1976).

5. APG IV, An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG IV. *Botanical Journal of the Linnean Society,* 181, 1-20 (2016).

6. J. M. Archibald, A. G. B. Simpson, & C. H. Slamovits, Eds. *Handbook of the Protists*. (Springer Nature, 2017).

7. N. Arisue, & T. Hashimoto, Phylogeny and evolution of apicoplasts and apicomplexan parasites. *Parasitology international*, 64, 254-259 (2015).

8. N. Arisue, *et al.*, Apicoplast phylogeny reveals the position of *Plasmodium vivax* basal to the Asian primate malaria parasite clade. *Scientific reports*, 9, 7274 (2019).

9. C. Azevedo, & P. M. Hine, Haplosporidia. In *Handbook of the Protists*, J. M. Archibald, A. G. B. Simpson, & C. H. Slamovits, Eds.(Springer Nature, 2017) pp.823-850.

10. C. Azevedo, & L. Corral, Some ultrastructural observations of a thraustochytrid (Protoctista, Labyrinthulomycota) from the clam *Ruditapes decussatus* (Mollusca, Bivalvia). *Diseases of aquatic organisms*, 31, 73-78 (1997).

11. D. Ballantine, & J. Norris, *Verdigellas*, a new deep water genus (Tetrasporales, Chlorophyta) from the tropical western Atlantic. *Cryptogamic Botany*, 4, 368-368 (1994).

12. R. N. Band, & C. Machemer, Environmental induction of multinucleate *Hartmannella rhysodes*. *Experimental Cell Research*, 31, 31-38 (1963).

13. D. Barthel, & A. Detmer, The spermatogenesis of *Halichondria panicea* (Porifera, Demospongiae). *Zoomorphology*, 110, 9-15 (1990).

14. D. Bass *et al.*, Clarifying the relationships between microsporidia and cryptomycota. Journal of Eukarotic Microbiology, 65, 773–782 (2018).

15. K. S. Bateman *et al.*, Single and multi-gene phylogeny of Hepatospora (Microsporidia)–a generalist pathogen of farmed and wild crustacean hosts. *Parasitology*, 143, 971-982 (2016).

16. G. W. Beakes & M. Thines, "Hyphochytriomycota and Oomycota". In *Handbook of the Protists*, J. M. Archibald, A. G. B. Simpson, & C. H. Slamovits, Eds. (Springer Nature, 2017) pp.435-506.

17. R. M. Bennett, D. Honda, G. W. Beakes, & M. Thines, "Labyrinthulomycota". In *Handbook of the Protists*, J. M. Archibald, A. G. B. Simpson, & C. H. Slamovits, Eds. (Springer Nature, 2017) pp.507-542.

18. R. Bernander, J. D. Palm, & S. G. Svärd, Genome ploidy in different stages of the *Giardia lamblia* life cycle. Cellular microbiology, 3, 55-62 (2001).

19. C. Berney *et al.*, Expansion of the 'Reticulosphere': Diversity of Novel Branching and Network-forming Amoebae Helps to Define Variosea (Amoebozoa). *Protist*, 166, 271-295 (2015).

20. R. Betancur *et al.*, The tree of life and a new classification of bony fishes. *PLoS currents*, 5 (2013).

21. S. K. Bhatnagar, Cytological perspective of Charophyta. II—*Nitella furcata* complex. *Phykos*, 28, 166–177 (1989).

22. D. Bhattacharya *et al.*, Genome of the red alga *Porphyridium purpureum*. Nature communications, 4, 1941 (2013).

23. C. W. Jr. Birky, J. Adams, M. Gemmel, & J. Perry, Using population genetic theory and DNA sequences for species detection and identification in asexual organisms. *PLoS one*, 5, 10609 (2010).

24. G. Bloomfield *et al.*, Triparental inheritance in *Dictyostelium*. *Proceedings of the National Academy of Sciences*, 116, 2187-2192 (2019).

25. D. Boltovskoy, O. R. Anderson, & N. Correa, "Radiolaria and Phaeodaria". In *Handbook of the Protists*, J. M. Archibald, A. G. B. Simpson, & C. H. Slamovits, Eds. (Springer Nature, 2017) pp.731-764.

26. M. W. Brown *et al.*, Phylogenomics demonstrates that breviate flagellates are related to opisthokonts and apusomonads. *Proceedings of the Royal Society of London B, Biological Sciences*, 280, 20131755 (2013).

27. G. Brugerolle, *Cryptophagus subtilis*: a new parasite of cryptophytes affiliated with the Perkinsozoa lineage. *European Journal of Protistology*, 37, 379-390 (2002).

28. S. Bulman, & S. Neuhauser, "Phytomyxea". In *Handbook of the Protists*, J. M. Archibald, A. G. B. Simpson, & C. H. Slamovits, Eds. (Springer Nature, 2017) pp.783-804.

29. F. Burki *et al.*, Untangling the early diversification of eukaryotes: A phylogenomic study of the evolutionary origins of Centrohelida, Haptophyta and Cryptista. *Proceedings of the Royal Society B, Biological Sciences*, 283, 20152802 (2016).

30. F. Burki, Evolution of Rhizaria: new insights from phylogenomic analysis of uncultivated protists. *BMC evolutionary biology*, 10, 377 (2010).

31. T. J. Byers, Growth, reproduction, and differentiation in Acanthamoeba. *International review of cytology*, 61, 283-338 (1979).

32. A. Cali, J. J. Becnel, & P. M. Takvorian, "Microsporidia". In *Handbook of the Protists*, J. M. Archibald, A. G. B. Simpson, & C. H. Slamovits, Eds. (Springer Nature, 2017) pp.1559-1618.

33. T. Cavalier-Smith, Kingdom protozoa and its 18 phyla. *Microbiology and Molecular Biology Reviews*, 57, 953-994 (1993).

34. T. Cavalier-Smith, Kingdom Chromista and its eight phyla: a new synthesis emphasising periplastid protein targeting, cytoskeletal and periplastid evolution, and ancient divergences. *Protoplasma*, 255, 297-357 (2018).

35. T. Cavalier-Smith, & J. M. Scoble, Phylogeny of Heterokonta: *Incisomonas marina*, a uniciliate gliding opalozoan related to *Solenicola* (Nanomonadea), and evidence that Actinophryida evolved from raphidophytes. *European Journal of Protistology*, 49, 328-353 (2013).

36. T. Cavalier-Smith, E. E. Chao, & R. Lewis, 187-gene phylogeny of protozoan phylum Amoebozoa reveals a new class (Cutosea) of deep-branching, ultrastructurally unique, enveloped marine Lobosa and clarifies amoeba evolution. *Molecular Phylogenetics and Evolution*, 99, 275–296 (2016).

37. T. Cavalier-Smith, E. E. Chao, & R. Lewis, Multigene phylogeny and cell evolution of chromist infrakingdom Rhizaria: contrasting cell organisation of sister phyla Cercozoa and Retaria. *Protoplasma*, 255, 1517-1574 (2018).

38. T. Cavalier-Smith *et al.*, Multigene phylogeny resolves deep branching of Amoebozoa. *Molecular Phylogenetics and Evolution*, 83, 293–304 (2015).

39. I. Čepička, M. F. Dolan, & G. H. Gile, "Parabasalia". In *Handbook of the Protists*, J. M. Archibald, A. G. B. Simpson, & C. H. Slamovits, Eds. (Springer Nature, 2017) pp.1175-1218.

40. L. R. Cleveland, The origin and evolution of meiosis. *Science*, 105, 287-289 (1947).

41. M. E. Cook, & L. E. Graham, "Chlorokybophyceae, Klebsormidiophyceae, Coleochaetophyceae". In *Handbook of the Protists*, J. M. Archibald, A. G. B. Simpson, & C. H. Slamovits, Eds. (Springer Nature, 2017) pp.185-204.

42. W. R. I. Cook, The life-history of *Cystochytrium radicale* occurring in the roots of *Veronica beccabunga*. *Transactions of the British Mycological Society*, 16, 246-252 (1932).

43. N. Corradi, Microsporidia: eukaryotic intracellular parasites shaped by gene loss and horizontal gene transfers. *Annual review of microbiology*, 69, 167-183 (2015).

44. E. W. Daniels, & G. D. Pappas, Reproduction of nuclei in *Pelomyxa palustris*. *Cell biology international*, 18, 805-812 (1994).

45. A. J. Dave, & M. B. Godward, Ultrastructural studies in the Rhodophyta. I. Development of mitotic spindle poles in *Apoglossum ruscifolium*, Kylin. *Journal of cell science*, 58, 345-362 (1982).

46. O. De Clerck, K. A. Bogaert, & F. Leliaert, Diversity and Evolution of Algae: Primary Endosymbiosis. *Advances in Botanical Research*, 64, 55-86 (2012).

47. C. P. De Souza, & S. A. Osmani, Mitosis, not just open or closed. *Eukaryotic cell*, 6, 1521-1527 (2007).

48. R. Derelle, P. López-García, H. Timpano, & D. Moreira, Phylogenomic Framework to Study the Diversity and Evolution of Stramenopiles (= Heterokonts). *Molecular Biology and Evolution*, 33, 2890–2898 (2016).

49. C. C. Dobell, The structure and life-history of *Copromonas subtilis*, nov. gen. et nov. spec.: a contribution to our knowledge of the Flagellata. *The Quarterly Journal of Microscopical Science*, 52, 75-120 (1908).

50. M. F. Dolan, U. D'Ambrosio, A. M. Wier, & L. Margulis, Surface kinetosomes and disconnected nuclei of a calonymphid: ultrastructure and evolutionary significance of *Snyderella tabogae*. *Acta Protozoologica*, 39, 135-142 (2000).

51. M. F. Dolan, A. M. Wier, & L. Margulis, Budding and asymmetric reproduction of a trichomonad with as many as 1000 nuclei in karyomastigonts: Metacoronympha from Incisitermes. *Acta Protozoologica*, 39, 275-280 (2000).

52. R. G. Dorrell, & C.J. Howe, Integration of plastids with their hosts: Lessons learned from dinoflagellates. *Proceedings of the National Academy of Sciences*, 112, 10247-10254 (2015).

53. W. Eikrem *et al.*, "Haptophyta". In *Handbook of the Protists*, J. M. Archibald, A. G. B. Simpson, & C. H. Slamovits, Eds. (Springer Nature, 2017) pp.893-954.

54. M. Eliáš *et al.*, "Eustigmatophyceae". In *Handbook of the Protists*, J. M. Archibald, A. G. B. Simpson, & C. H. Slamovits, Eds. (Springer Nature, 2017) Cham, pp.367-406.

55. H. Ettl, Xanthophyceae. In *Susswasserflora von Mitteleuropa (Bd. 3, I. Teil), H.* Ettl, H.J. Gerloff, & H. Heynig, Eds. (Gustav Fischer, 1987).

56. R. I. Figueroa, M. Estrada, & E. Garcés, Life histories of microalgal species causing harmful blooms: Haploids, diploids and the relevance of benthic stages. *Harmful algae*, 73, 44-57 (2018).

57. K. Finstermeier *et al.*, A mitogenomic phylogeny of living primates. *PLoS One*, 8, e6950 (2013).

58. A. J. Flemming, Z. Z. Shen, A. Cunha, S. W. Emmons, & A. M. Leroi, Somatic polyploidization and cellular proliferation drive body size evolution in nematodes. *Proceedings of the National Academy of Scien*ces, 97, 5285-5290 (2000).

59. J. Fraga *et al.*, Phylogeny of *Leishmania* species based on the heat-shock protein 70 gene. *Infection, Genetics and Evolution*, 10, 238-245 (2010).

60. W. Franzen, Oogenesis and larval development of *Scypha ciliata* (Porifera, Calcarea). *Zoomorphology*, 107, 349-357 (1988).

61. S.G. Garg, & M. F. Martin, Mitochondria, the cell cycle, and the origin of sex via a syncytial eukaryote common ancestor. *Genome biology and evolution*, 8, 1950-1970 (2016).

62. W. Gibson, "Kinetoplastea". In *Handbook of the Protists*, J. M. Archibald, A. G. B. Simpson, & C. H. Slamovits, Eds. (Springer Nature, 2017) pp.1089-1138.

63. M.C. Grant, & V. W. Proctor, *Chara vulgaris* and *C. contraria*: patterns of reproductive isolation for two cosmopolitan species complexes. *Evolution*, 267-281 (1972).

64. M. C. Grant, & V. W. Proctor, Electrophoretic analysis of genetic variation in the Charophyta. I. Gene duplication via polyploidy. *Journal of Phycology*, 16, 109-115 (1980).

65. L. Guillou *et al.*, Widespread occurrence and genetic diversity of marine parasitoids belonging to Syndiniales (Alveolata). *Environmental microbiology*, 10, 3349-336 (2008).

66. D. Haig, Retroviruses and the placenta. *Current Biology*, 22, R609-R613 (2012).

67. J. D. Hall, & R. M. McCourt, "Zygnematophyta". In *Handbook of the Protists*, J. M. Archibald, A. G. B. Simpson, & C. H. Slamovits, Eds. (Springer Nature, 2017) pp.135-164.

68. V. Hampl, "Preaxostyla". In *Handbook of the Protists*, J. M. Archibald, A. G. B. Simpson, & C. H. Slamovits, Eds. (Springer Nature, 2017) pp.1139-1174.

69. V. Hartenstein, *Atlas of* Drosophila *development.* (Cold Spring Harbor Laboratory Press, 2013).

70. K. Hausmann, N. Hülsmann, & R. Radek, Protistology. In *E. Schweizerbart'sche Verlagsbuchhandlung*, (Stuttgart, 1993).

71. D. He, O. Fiz-Palacios, C. J. Fu, J. Fehling, C. C. Tsai, & S. L. Baldauf, An alternative root for the eukaryote tree of life. *Current Biology*, 24, 465-470 (2014).

72. A. A. Heiss, M. W. Brown, & G. B. Simpson, "Apusomonadida". In *Handbook of the Protists*, J. M. Archibald, A. G. B. Simpson, & C. H. Slamovits, Eds. (Springer Nature, 2017) pp.1619-1646.

73. K. M. Helgen, The mammal family tree. *Science*, 334, 458-459 (2011).

74. M. L. Hernandez-Nicaise, G. Nicaise, & L. Malaval, Giant Smooth Muscle Fibers of the *Ctenophore Mnemiopsis leydii*: Ultrastructural Study of in situ and Isolated Cells. *The Biological Bulletin*, 167, 210-228 (1984).

75. S. Hess, N. Sausen, & M. Melkonian, Shedding light on vampires: the phylogeny of vampyrellid amoebae revisited. *PLoS One*, 7, e31165 (2012).

76. M. T. Higham, & T. Bisalputra, A further note on the surface structure of *Scenedesmus coenobium*. *Canadian journal of botany*, 48, 1839-1841 (1970).

77. Y. Hirakawa, & K. I. Ishida, Polyploidy of endosymbiotically derived genomes in complex algae. *Genome biology and evolution*, 6, 974-980 (2014).

78. K. Hoef-Emden, & J. M. Archibald, "Cryptophyta (Cryptomonads)". In *Handbook of the Protists*, J. M. Archibald, A. G. B. Simpson, & C. H. Slamovits, Eds. (Springer Nature, 2017) pp.851-891

79. P.G. Hofstatter, & D.J. Lahr, All Eukaryotes Are Sexual, unless Proven Otherwise: Many So-Called Asexuals Present Meiotic Machinery and Might Be Able to Have Sex. *BioEssays*, 1800246 (2019).

80. S. Hongsanan *et al.*, An updated phylogeny of Sordariomycetes based on phylogenetic and molecular clock evidence. *Fungal diversity*, 84, 25-41 (2017).

81. T. Horiguchi, "Raphidophyceae (Raphidophyta)". In *Handbook of the Protists*, J. M. Archibald, A. G. B. Simpson, & C. H. Slamovits, Eds. (Springer, 2017) pp.305-330.

82. S. Horn *et al.*, *Synchroma grande* spec. nov. (Synchromophyceae class. nov., Heterokontophyta): an amoeboid marine alga with unique plastid complexes. *Protist*, 158, 277-293 (2007).

83. L. C. Hughes *et al.*, Comprehensive phylogeny of ray-finned fishes (Actinopterygii) based on transcriptomic and genomic data. *Proceedings of the National Academy of Sciences*, 115, 6249-6254 (2018).

84. N. A. Irwin *et al.*, Phylogenomics supports the monophyly of the Cercozoa. *Molecular phylogenetics and evolution*, 130, 416-423 (2019).

85. C. Jackson, S. Clayden, & A. Reyes-Prieto, The Glaucophyta: the blue-green plants in a nutshell. *Acta Societatis Botanicorum Poloniae*, 84, 149-165 (2015).

86. T. Y. James, T. M. Porter, & W. W. Martin, Blastocladiomycota. In *Systematics and Evolution: The Mycota VII part A,* D. J. McLaughlin, & J. W. Spatafora, Eds. (Springer, 2014)

87. J. Janouškovec *et. al.*, Factors mediating plastid dependency and the origins of parasitism in apicomplexans and their close relatives. *Proceedings of the National Academy of Sciences of the United States of America*, 112, 10200–10207 (2015).

88. R. Kamikawa *et al*, Gene content evolution in discobid mitochondria deduced from the phylogenetic position and complete mitochondrial genome of *Tsukubamonas globosa*. *Genome Biolology and Evolution*, 6, 306-315 (2014).

89. S. A. Karpov *et al.*, Morphology, phylogeny, and ecology of the aphelids (Aphelidea, Opisthokonta) and proposal for the new superphylum Opisthosporidia. *Frontiers in Microbiology*, 5, 112 (2014).

90. L. A. Katz, & J. R. Grant, Taxon-rich phylogenomic analyses resolve the eukaryotic tree of life and reveal the power of subsampling by sites. *Systematic biology*, 64, 406-415 (2014).

91. H. Kawai, & E. C. Henry, "Phaeophyta". In *Handbook of the Protists*, J. M. Archibald, A. G. B. Simpson, & C. H. Slamovits, Eds. (Springer Nature, 2017) pp.267-304.

92. C. Koch *et al.*, The life cycle of the amoeboid alga *Synchroma grande* (Synchromophyceae, Heterokontophyta)–highly adapted yet equally equipped for rapid diversification in benthic habitats. *Plant Biology*, 13, 801-808 (2011).

93. M. Kostka, "Opalinata". In *Handbook of the Protists*, J. M. Archibald, A. G. B. Simpson, & C. H. Slamovits, Eds. (Springer Nature, 2017) pp.543-566.

94. A. K. Krabberød *et al.*, Single cell transcriptomics, mega-phylogeny, and the genetic basis of morphological innovations in Rhizaria. *Molecular Biology and Evolution*, 34, 1557-1573 (2017).

95. J. Kristiansen, & P. Škaloud, "Chrysophyta". In *Handbook of the Protists*, J. M. Archibald, A. G. B. Simpson, & C. H. Slamovits, Eds. (Springer Nature, 2017) pp.331-366.

96. J. Kulda, E. Nohýnková, & I. Čepička, "Retortamonadida" (with Notes on Carpediemonas-Like Organisms and Caviomonadidae). In *Handbook of the Protists*, J. M. Archibald, A. G. B. Simpson, & C. H. Slamovits, Eds. (Springer Nature, 2017) pp.1247-1278.

97. C. E. Laumer *et al.*, Spiralian phylogeny informs the evolution of microscopic lineages. *Current Biology*, 25, 2000-2006 (2015).

98. B. S. Leander, G. Lax, A. Karnkowska, & A. G. B. Simpson, "Euglenida". In *Handbook of the Protists*, J. M. Archibald, A. G. B. Simpson, & C. H. Slamovits, Eds. (Springer Nature, 2017) pp.1047-1088.

99. F. Leliaert *et al.*, Phylogeny and molecular evolution of the green algae. *Critical Reviews in Plant Sciences*, 31, 1–46 (2012).

100. F. Leliaert, H. Verbruggen, & F. W. Zechman, Into the deep: New discoveries at the base of the green plant phylogeny. *BioEssays*, 33, 683–692 (2011).

101. C. Lemieux, C. Otis, & M. Turmel, Chloroplast phylogenomic analysis resolves deep-level relationships within the green algal class Trebouxiophyceae. *BMC evolutionary biology*, 14, 211 (2014).

102. R. J. G. Lester, & P. M. Hine, "Paramyxida". In *Handbook of the Protists*, J. M. Archibald, A. G. B. Simpson, & C. H. Slamovits, Eds. (Springer Nature, 2017) pp.805-822.

103. S. K. Maciver, Asexual amoebae escape Muller's ratchet through polyploidy. *Trends in parasitology*, 32, 855-862 (2016).

104. S. Maistro, P. Broady, C. Andreoli, & E. Negrisolo, "Xanthophyceae". In *Handbook of the Protists*, J. M. Archibald, A. G. B. Simpson, & C. H. Slamovits, Eds. (Springer Nature, 2017) pp.407-434.

105. D. G. Mann, R. M. Crawford, & F. E. Round, "Bacillariophyta". In *Handbook of the Protists*, J. M. Archibald, A. G. B. Simpson, & C. H. Slamovits, Eds. (Springer Nature, 2017) pp.205-266.

106. H. J. Marchant, & J. D. Pickett-Heaps, Ultrastructure and differentiation of *Hydrodictyon reticulatum* I. Mitosis in the coenobium. *Australian journal of biological sciences*, 23, 1173-1186 (1970).

107. Marcili, A. *et al.*, Phylogenetic relationships of *Leishmania* species based on trypanosomatid barcode (SSU rDNA) and gGAPDH genes: Taxonomic revision of *Leishmania (L.) infantum chagasi* in South America. *Infection, Genetics and Evolution*, 25, 44-51 (2014).

108. W. L. Marshall, G. Celio, D. J. McLaughlin, & M. L. Berbee, Multiple isolations of a culturable, motile Ichthyosporean (Mesomycetozoa, Opisthokonta), *Creolimax fragrantissima* n. gen., n. sp., from marine invertebrate digestive tracts. *Protist*, 159, 415-433 (2008).

109. C. G. McCarthy, & D. A. Fitzpatrick, Multiple approaches to phylogenomic reconstruction of the fungal kingdom. *Advances in genetics, Academic Press*, 100, 211-266 (2017).

110. R. M. McCourt, K. G. Karol, J. D. Hall, M. T. Casanova, & M. C. Grant, "Charophyceae (Charales)". In *Handbook of the Protists*, J. M. Archibald, A. G. B. Simpson, & C. H. Slamovits, Eds. (Springer Nature, 2017) pp.165-184.

111. R. W. Meredith *et al.*, Impacts of the Cretaceous Terrestrial Revolution and KPg extinction on mammal diversification. *Science*, 334, 521-524 (2011).

112. J. Mignot, Study of the cell nucleus of the euglenoid *Scytomonas pusilla* (Stein) during division and copulation. *Comptes rendus hebdomadaires des seances de l'Academie des sciences*, 254, 1864-1866 (1962).

113. B. Misof, Phylogenomics resolves the timing and pattern of insect evolution. *Science*, 346, 763-767 (2014).

114. M. Müller *et al.*, Biochemistry and evolution of anaerobic energy metabolism in eukaryotes. *Microbiology and Molecular Biology Reviews*, 76, 444-495 (2012).

115. C. Nagasato, & T. Motomura, Ultrastructural study on mitosis and cytokinesis in *Scytosiphon lomentaria* zygotes (Scytosiphonales, Phaeophyceae) by freeze-substitution. *Protoplasma*, 219, 140-149 (2002).

116. H. W. Nichols, Culture and developmental morphology of *Compsopogon coeruleus*. *American Journal of Botany*, 51, 180-188 (1964).

117. C. Nielsen, *Animal evolution: interrelationships of the living phyla*. (Oxford, 2012).

118. K. J. Niklas, E. D. Cobb, & D. R. Crawford, The evo-devo of multinucleate cells, tissues, and organisms, and an alternative route to multicellularity. *Evolution & development*, 15, 466-474 (2013).

119. S. I. Nikolaev *et al.*, The twilight of Heliozoa and rise of Rhizaria, an emerging supergroup of amoeboid eukaryotes. *Proceedings of the National Academy of Sciences*, 101, 8066–8071 (2004).

120. P. M. O'Grady, & R. DeSalle, Phylogeny of the genus *Drosophila*. *Genetics*, 209, 1-25 (2018).

121. M. Oborník *et al.*, Morphology, ultrastructure and life cycle of *Vitrella brassicaformis* n. sp., n. gen., a novel chromerid from the Great Barrier Reef. *Protist*, 163, 306-323 (2012).

122. J. L. Olefeld, S. Majda, D. C. Albach, S. Marks, & J. Boenigk, Genome size of chrysophytes varies with cell size and nutritional mode. *Organisms Diversity & Evolution*, 18, 163-173 (2018).

123. L. Oliveira, & T. Bisalputra, Ultrastructural and cytochemical studies on the nature and origin of the cytoplasmic inclusions of aging cells of Ectocarpus (Phaeophyta, Ectocarpales). *Phycologia*, 16, 235-243 (1977).

124. S. Ota, A. Kudo, K. I. & Ishida, *Gymnochlora dimorpha* sp. nov., a chlorarachniophyte with unique daughter cell behaviour. *Phycologia*, 50, 317-326 (2011).

125. T. Pánek, A. G. B. Simpson, M. W. Brown, & B. D. Dyer, "Heterolobosea".In *Handbook of the Protists*, J. M. Archibald, A. G. B. Simpson, & C. H. Slamovits, Eds. (Springer Nature, 2017) pp.1005-1046.

126. J. S. Park, & A. G. B. Simpson, Diversity of heterotrophic protists from extremely hypersaline habitats. *Protist*, 166, 422–437 (2015).

127. P. Perelman *et al.*, A molecular phylogeny of living primates. *PLoS genetics*, 7, e1001342 (2011).

128. R. S. Peters *et al.*, Evolutionary history of the Hymenoptera. *Current Biology*, 27, 1013-1018 (2017).

129. G. Piganeau, N. Grimsley, & H. Moreau, Genome diversity in the smallest marine photosynthetic eukaryotes. *Research in microbiology*, 162, 570-577 (2011).

130. M. J. Powell, "Blastocladiomycota". In *Handbook of the Protists*, J. M. Archibald, A. G. B. Simpson, & C. H. Slamovits, Eds. (Springer Nature, 2017a) pp.1497-1522.

131. M. J. Powell, "Chytridiomycota". In *Handbook of the Protists*, J. M. Archibald, A. G. B. Simpson, & C. H. Slamovits, Eds. (Springer Nature, 2017b) pp.1523-1558.

132. M. J. Powell, & P. M. Letcher, "6 Chytridiomycota, Monoblepharidomycota, and Neocallimastigomycota". In *The Mycota Part VII A. Systematics and Evolution, 2nd Edition*, D. J. McLaughlin, & J. W. Spatafora, Eds. (Springer, 2014), pp. 141-175.

133. M. K. Poxleitner *et al*., Evidence for karyogamy and exchange of genetic material in the binucleate intestinal parasite *Giardia intestinalis*. *Science*, 319, 1530-1533 (2008).

134. D. C. Price *et al*., "Glaucophyta". In *Handbook of the Protists*, J. M. Archibald, A. G. B. Simpson, & C. H. Slamovits, Eds. (Springer Nature, 2017) pp.23-88.

135. R. O. Prum *et al*., A comprehensive phylogeny of birds (Aves) using targeted next-generation DNA sequencing. *Nature*, 526, 569 (2015).

136. M. Pussard, Comparaison morphologique de 4 souches d'Acanthamoeba du groupe astronyxis-comandoni. *The Journal of Protozoology*, 19, 557-563 (1972).

137. S. Raghu-Kumar, Occurrence of the thraustochytrid, *Corallochytrium limacisporum* gen. et sp. nov. in the coral reef lagoons of the Lakshadweep Islands in the Arabian Sea. *Botanica Marina*, 30, 83-90 (1987).

138. I. B. Raikov, The protozoan nucleus: Morphology and Evolution. *Cell biology monographs*, 9, 266-364 (1982).

139. I. B. Raikov, The diversity of forms of mitosis in protozoa: a comparative review. *European Journal of Protistology*, 30, 253-269 (1994).

140. J. A. Raven, F. A. Smith, & S.M. Glidewell, Photosynthetic capacities and biological strategies of giant-celled and small-celled macro-algae. *New Phytologist*, 83, 299-309 (1979).

141. A. Reñé, E. Alacid, I. Ferrera, & E. Garcés, Evolutionary Trends of Perkinsozoa (Alveolata) Characters Based on Observations of Two New Genera of Parasitoids of dinoflagellates, *Dinovorax* gen. nov. and *Snorkelia* gen. nov.. *Frontiers in microbiology*, 8, 1-16 (2017).

142. I. Riisberg *et al*., Seven gene phylogeny of heterokonts. *Protist*, 160, 191–204 (2009).

143. G. Röderer, Hemmung der Cytokinese und Bildung von Riesenzellen bei *Poterioochromonas malhamensis* durch organische Bleiverbindungen und andere Agenzien. *Protoplasma*, 99, 39-51 (1979).

144. A. J. Roger, S. A. Muñoz-Gómez, & R. Kamikawa, The origin and diversification of mitochondria. *Current Biology*, 27, R1177-R1192 (2017).

145. T. G. Rosser *et. al*., Arrested Development of *Henneguya ictaluri* (Cnidaria: Myxobolidae) in ♀ Channel Catfish × ♂ Blue Catfish Hybrids. *Journal of Aquatic Animal Health*, 31, 201-213 (2019).

146. J. F. Ryan, & M. Chiodin, Where is my mind? How sponges and placozoans may have lost neural cell types. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 370, 20150059 (2015).

147. U. Saller, Oogenesis and larval development of *Ephydatia fluviatilis* (Porifera, Spongillidae). *Zoomorphology*, 108, 23-28 (1988).

148. N. Salmaso, & M. Tolotti, "Other phytoflagellates and groups of lesser importance". In *Encyclopedia of inland waters*, G. E. Likens, Ed. (Elsevier, 2009) pp. 174-183.

149. W. B. Sanders, R. L. Moe, & C. Ascaso, Ultrastructural study of the brown alga *Petroderma maculiforme* (Phaeophyceae) in the free- living state and in lichen symbiosis with the intertidal marine fungus *Verrucaria tavaresiae* (Ascomycotina). *European Journal of Phycology*, 40, 353-361 (2005).

150. A. C. Schmit, & P. Nick, "Microtubules and the evolution of mitosis". In Plant Microtubules. Plant Cell Monographs, P. Nick, Ed. (Springer, Heidelberg, 2008) pp. 233-266.

151. G. Schönian, J. Lukeš, O. Stark, & J. A. Cotton, "Molecular evolution and phylogeny of *Leishmania"*. In Drug Resistance in Leishmania Parasites, A. Ponte-Sucre, & M. Padrón-Nieves (Springer, 2010) pp. 19-57.

152. G. Schönian, I. Mauricio, & E. Cupolillo, Is it time to revise the nomenclature of *Leishmania*?. *Trends in parasitology*, 26, 466-469 (2010).

153. C. Scornavacca, & N. Galtier, Incomplete lineage sorting in mammalian phylogenomics. *Systematic biology*, 66, 112-120 (2017).

154. J. Scott Mitosis in the freshwater red alga *Batrachospermum octocarpum*. *Protoplasma*, 118, 56 -70 (1983).

155. A. Sebé-Pedrós *et al*., Regulated aggregative multicellularity in a close unicellular relative of metazoa. *Elife*, 2, 1-20 (2013).

156. T. Shiratori, T. Nakayama, & K. Ishida, A new deep-branching stramenopile, *Platysulcus tardus* gen. nov., sp. nov.. *Protist*, 166, 337–348 (2015).

157. R. Sierra et al., Evolutionary origins of Rhizarian parasites. *Molecular Biology and Evolution*, 33, 980-983 (2016).

158. R. Sierra *et al*., Deep relationships of Rhizaria revealed by phylogenomics: a farewell to Haeckel's Radiolaria. *Molecular Phylogenetics and Evolution*, 67, 53–59 (2013).

159. A. G. B. Simpson, "Jakobida". In *Handbook of the Protists*, J. M. Archibald, A. G. B. Simpson, & C. H. Slamovits, Eds. (Springer Nature, 2017) pp.973-1004.

160. A. G. B. Simpson, C. H. Slamovits, & J. M. Archibald, Protist Diversity and Eukaryote Phylogeny. In *Handbook of the Protists*, J. M. Archibald, A. G. B. Simpson, & C. H. Slamovits, Eds. (Springer Nature, 2017) pp.1-21.

161. R. J. Soreng *et al*., A worldwide phylogenetic classification of the Poaceae (Gramineae) II: An update and a comparison of two 2015 classifications. *Journal of Systematics and Evolution*, 55, 259-290 (2017).

162. J. W. Spatafora, The Fungal Tree of Life: from Molecular Systematics to Genome-Scale Phylogenies. *ASM Science, Microbiology Spectrum*, 1-32 (2017).

163. J. W. Spatofora, A phylum-level phylogenetic classification of zygomycete fungi based on genome-scale data. *Mycologia*, 108, 1028-1046 (2016).

164. D. Speijer, J. Lukeš, & M. Eliáš, Sex is a ubiquitous, ancient, and inherent attribute of eukaryotic life. *Proceedings of the National Academy of Sciences*, 112, 8827-8834 (2015).

165. V. Sprague, *Minchinia louisiana* n. sp.(Haplosporidia, Haplosporidiidae), a parasite of Panopeus herbstii. *The Journal of Protozoology*, 10, 267-274 (1963).

166. M. Srivastava *et al*., The *Trichoplax* genome and the nature of placozoans. *Nature*, 454, 955–960 (2008).

167. A. Stechmann, & T. Cavalier-Smith, Rooting the eukaryote tree by using a derived gene fusion. *Science*, 297, 89-91 (2002).

168. J. M. Steiner, Technical notes: Growth of *Cyanophora paradoxa*. *Journal of Endocytobiosis & Cell Research*, 20, 62-67 (2010).

169. E. C. Swart *et al.*, The *Oxytricha trifallax* macronuclear genome: a complex eukaryotic genome with 16,000 tiny chromosomes. *PLoS biology*, 11, e1001473 (2013).

170. C. Székely *et al.*, A synopsis of records of myxozoan parasites (Cnidaria: Myxozoa) from shrews, with additional data on *Soricimyxum fegati* from common shrew *Sorex araneus* in Hungary and pygmy shrew *Sorex minutus* in Slovakia. *Folia parasitologica*, 63, 1-5 (2016).

171. G. Tanifuji *et al.*, The draft genome of *Kipferlia bialata* reveals reductive genome evolution in fornicate parasites. *PloS one*, 13, e0194487 (2018).

172. L. Tedersoo *et al.*, High-level classification of the Fungi and a tool for evolutionary ecological analyses. *Fungal Diversity*, 90, 135-159 (2018).

173. A. K. Tice *et al.*, *Sorodiplophrys stercorea*: another novel lineage of sorocarpic multicellularity. *Journal of Eukaryotic Microbiology*, 63, 623–628 (2016).

174. D. V. Tikhonenkov *et al.*, Description of *Colponema vietnamica* sp. n. and *Acavomonas peruviana* n. gen. n. sp., two new Alveo- late phyla (Colponemidia nom. nov. and Acavomonidia nom. nov.) and their contributions to reconstructing the ancestral state of alveolates and eukaryotes. *PLoS ONE*, 9, e95467 (2014).

175. G. Torruella *et al.*, Phylogenomics reveals convergent evolution of lifestyles in close relatives of animals and fungi. *Current Biology*, 25, 2404–2410 (2015).

176. H. Trenkwalder, Neue Bodenalgen aus Föhrenwäldern im Raum von Brixen (Südtirol, Italien). *Berichte des Naturwissenschaftlich-medizinischen Vereins in Innsbruck*, 62, 7-19 (1975).

177. M. Turmel, 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*, 26, 631-648 (2008).

178. S. Tyler, & M. Hooge, Comparative morphology of the body wall in flatworms (Platyhelminthes). *Canadian Journal of Zoology*, 82, 194-210 (2004).

179. C. van den Hoek, D. G. Mann, & H. M. Jahns, *Algae: An introduction of Phycology*. (Cambridge University Press, 1995).

180. P. Vďačný *et al.*, Morphological and molecular phylogeny of dileptid and tracheliid ciliates: Resolution at the base of the class Litostomatea (Ciliophora, Rhynchostomatia). *European journal of protistology*, 47, 295-313 (2011).

181. P. Von Aderkas *et al.*, Multinucleate storage cells in Douglas-fir (*Pseudotsuga menziesii* (Mirbel) Franco) and the effect of seed parasitism by the chalcid *Megastigmus spermotrophus* Wachtl. Heredity, 94, 616 (2005).

182. C. R. Vossbrinck et al., Ribosomal RNA sequence suggests microsporidia are extremely ancient eukaryotes. *Nature*, 326, 411 (1987).

183. G. Walker, J. B. Dacks, & T. Martin Embley, Ultrastructural description of *Breviata anathema*, n. gen., n. sp., the organism previously studied as "*Mastigamoeba invertens*". *Journal of Eukaryotic Microbiology*, 53, 65-78 (2006).

184. S. Williams, XXIX.—A Contribution to the Experimental Morphology of *Lycopodium Selago*, with Special Reference to the Development of Adventitious Shoots. *Earth and Environmental Science Transactions of The Royal Society of Edinburgh*, 57, 711-737 (1934).

185. K. Willis, & J. McElwain, *The evolution of plants*. (Oxford University Press, 2014).

186. N. B. Willumsen, F. Siemensma, & P. Suhr-Jessen, A Multinucleate Amoeba, *Parachaos zoochlorellae* (Willumsen 1982) comb, nov., and a Proposed Division of the Genus Chaos into the Genera Chaos and Parachaos (Gymnamoebia, Amoebidae). *Archiv für Protistenkunde*, 134, 303-313 (1987).

187. A. Z. Worden, & F. Not, Ecology and diversity of picoeukaryotes. *Microbial ecology of the Oceans*, 2, 159-205 (2008).

188. E. C. Yang *et al.*, Divergence time estimates and the evolution of major lineages in the florideophyte red algae. *Scientific Reports*, 6, 21361 (2016).

189. H. S. Yoon *et al.*, Defining the major lineages of red algae (Rhodophyta). *Journal of Phycology*, 42, 482–492 (2006).

190. H. S. Yoon, "Rhodophyta". In *Handbook of the Protists*, J. M. Archibald, A. G. B. Simpson, & C. H. Slamovits, Eds. (Springer Nature, 2017) pp.89-134,

191. H. S. Yoon, G. C. Zuccarello, & D. Bhattacharya, Evolutionary history and taxonomy of red algae. In "*Red algae in genomic age*", J. Seckbach, & D. J. Chapman, Eds. (Springer, 2010) pp. 27-45.

Supplementary Table 4: List of all eukaryotic trees used in the ancestral state reconstruction. First column represents different root positions. A – Archaeplastida, E – Excavata, H – Hacrobia, M – Amoebozoa, O – Opisthokonta or Obazoa, and S – SAR; monophyly (holophyly) of each supergroup in the analysis is indicated by(+) if the group is monophyletic, (-) if not. Full Newick trees are shown for all 106 taxa.

ROOT POSITION	CODE	SHORT NEWICK	A	E	H	M	O	S	FULL NEWICK TREE OF 106 TAXA
AMORPHAEA (OPISTHOKONTA + AMOEBOZOA)	dicho_01	((M,O),(E,((A,H),S)));	+	+	+	+	+	+	((((Discosea,Tubulinea),((Archamoebea,Cutosea),(Variosea,(Myxogastria,(Protosporangiida,Dictyostelia))))),((Apusomonadida,Breviatea),((Rotosphaerida,(Microsporidia,(Blastocladiomycota,((Chytridiomycota,Neocallimastiqomycota),(Zoopagomycota,(Mucoromycota,(Basidiomycota,Ascomycota)))))),((Coralliochytrium,Mesomycetozoea),(Filasterea,(Choanoflagellata,((Demospongiae,Hexactinellida),((Calcarea,Homoscleromorpha),(Placozoa,(Ctenophora,(Cnidaria,(Deuterostomia,Proterostomia)))))))))),((((Preaxostyla,Parabasalia,(Calvomonadidae,(Chilomastix,(Kipferlia,(Dysnectes,(Diplomonadida,Retortamonas)))))),(Jakobida,(Heterolobosea,(Euglenida,(Kinetoplastea,Diplonemida))))),((((Cyanidiophyceae,((Porphyridiophyceae,(Compsopogonophyceae,Stylonematophyceae)),(Rhodellophyceae,(Florideophyceae,Bangiophyceae)))),Glaucophyta),((((Mesostigma,Chlorokybophyceae),(Klebsormidiophyceae,(Charophyceae,(Zygnematophyta,(Coleochaetophyceae,(Bryophyta,(Marchantiophyta,(Anthoceratophyta,(Euphyllophytina,Lycopodiophyta)))))))),((Filosa,(Paramyxida,Haplosporidia),(Phytomyxea,Vampyrellidae)),(Foraminifera,(Polycystinea,Acantharea)))),((Ciliophora,((Dinoflagellata,Perkinsozoa),(Apicomplexa,Chromerida))),(Labyrinthulomycota,(Blastocystida,Opalinata),(Oomycota,Hyphochytriomycota),((Pelagophyceae,Bacillariophyceae),(Actinophrida,((Raphidophyceae,(Xanthophyceae,Phaeophyceae)),(Chrysophyta,Eustigmatophyceae,Synchromophyceae))))))));
OBAZOA (OR OPISTHOKONTA)	dicho_02	(O,(M,(E,(S,(A,H)))));	+	+	+	+	+	+	(((Rotosphaerida,(Microsporidia,(Blastocladiomycota,((Chytridiomycota,Neocallimastiqomycota),(Zoopagomycota,(Mucoromycota,(Basidiomycota,Ascomycota)))))),((Coralliochytrium,Mesomycetozoea),(Filasterea,(Choanoflagellata,((Demospongiae,Hexactinellida),((Calcarea,Homoscleromorpha),(Placozoa,(Ctenophora,(Cnidaria,(Deuterostomia,Proterostomia)))))))))),((Apusomonadida,Breviatea),((Discosea,Tubulinea),(Archamoebea,Cutosea),(Variosea,(Myxogastria,(Protosporangiida,Dictyostelia))))),((((Preaxostyla,Parabasalia,(Calvomonadidae,(Chilomastix,(Kipferlia,(Dysnectes,(Diplomonadida,Retortamonas)))))),(Jakobida,(Heterolobosea,(Euglenida,(Kinetoplastea,Diplonemida))))),((((Cyanidiophyceae,((Porphyridiophyceae,(Compsopogonophyceae,Stylonematophyceae)),(Rhodellophyceae,(Florideophyceae,Bangiophyceae)))),Glaucophyta),((((Mesostigma,Chlorokybophyceae),(Klebsormidiophyceae,(Charophyceae,(Zygnematophyta,(Coleochaetophyceae,(Bryophyta,(Marchantiophyta,(Anthoceratophyta,(Euphyllophytina,Lycopodiophyta)))))))),((Filosa,(Paramyxida,Haplosporidia),(Phytomyxea,Vampyrellidae)),(Foraminifera,(Polycystinea,Acantharea)))),((Ciliophora,((Dinoflagellata,Perkinsozoa),(Apicomplexa,Chromerida))),(Labyrinthulomycota,(Blastocystida,Opalinata),(Oomycota,Hyphochytriomycota),((Pelagophyceae,Bacillariophyceae),(Actinophrida,((Raphidophyceae,(Xanthophyceae,Phaeophyceae)),(Chrysophyta,Eustigmatophyceae,Synchromophyceae))))))));
	poly_02	(O,(M,(E1,E2,(S,A,H1,H2))));	+	-	-	+	+	+	(((Rotosphaerida,(Microsporidia,(Blastocladiomycota,((Neocallimastiqomycota,Chytridiomycota),(Zoopagomycota,(Mucoromycota,(Basidiomycota,Ascomycota)))))),((Mesomycetozoea,Coralliochytrium),(Filasterea,(Choanoflagellata,(Homoscleromorpha,Calcarea,(Hexactinellida,Demospongiae),(Placozoa,(Cnidaria,(Deuterostomia,Proterostomia)))))))),((Breviatea,Apusomonadida),((Tubulinea,Discosea),(Variosea,Cutosea,Archamoebea,(Myxogastria,Protosporangiida,Dictyostelia)))),((Jakobida,(Heterolobosea,(Euglenida,(Diplonemida,Kinetoplastea)))),(Preaxostyla,(Parabasalia,(Caviomonadidae,(Chilomastix,(Kipferlia,(Dysnectes,(Diplomonadida,Retortamonas)))))))),((((Katablepharidea,(Cryptophyceae,Goniomonadida)),(Pavlovales,(Phaeocystales,(Prymnesiales,(Coccolithales,Isochrysidales))))),(Glaucophyta,(Cyanidiophyceae,((Rhodellophyceae,(Florideophyceae,Bangiophyceae)),(Porphyridiophyceae,(Compsopogonophyceae,Stylonematophyceae))))),(((Palmophyllales,Prasinococcales),(Mamiellales,Monomastigales),(Trebouxiophyceae,(Ulvophyceae,Chlorophyceae)))),((Mesostigma,Chlorokybophyceae),(Klebsormidiophyceae,(Charophyceae,(Zygnematophyta,(Coleochaetophyceae,(Bryophyta,(Marchantiophyta,(Anthoceratophyta,(Euphyllophytina,Lycopodiophyta)))))))),((Filosa,(Paramyxida,Haplosporidia),(Phytomyxea,Vampyrellidae)),(Foraminifera,(Polycystinea,Acantharea)))),((Ciliophora,((Dinoflagellata,Perkinsozoa),(Apicomplexa,Chromerida))),(Labyrinthulomycota,(Blastocystida,Opalinata),(Oomycota,Hyphochytriomycota),((Pelagophyceae,Bacillariophyceae),(Actinophrida,((Raphidophyceae,(Xanthophyceae,Phaeophyceae)),(Chrysophyta,Eustigmatophyceae,Synchromophyceae))))))));
AMOEBOZOA	dicho_03	(M,(O,(E,(S,(A,H)))));	+	+	+	+	+	+	((((Tubulinea,Discosea),((Cutosea,Archamoebea),(Variosea,(Myxogastria,(Protosporangiida,Dictyostelia))))),((Breviatea,Apusomonadida),((Rotosphaerida,(Microsporidia,(Blastocladiomycota,((Neocallimastiqomycota,Chytridiomycota),(Zoopagomycota,(Mucoromycota,(Basidiomycota,Ascomycota)))))),((Mesomycetozoea,Coralliochytrium),(Filasterea,(Choanoflagellata,((Hexactinellida,Demospongiae),((Homoscleromorpha,Calcarea),(Placozoa,(Ctenophora,(Cnidaria,(Deuterostomia,Proterostomia)))))))))),((Jakobida,(Heterolobosea,(Euglenida,(Diplonemida,Kinetoplastea)))),(Preaxostyla,(Parabasalia,(Caviomonadidae,(Chilomastix,(Kipferlia,(Dysnectes,(Diplomonadida,Retortamonas)))))))),((((Katablepharidea,(Cryptophyceae,Goniomonadida)),(Pavlovales,(Phaeocystales,(Prymnesiales,(Coccolithales,Isochrysidales))))),(Cyanidiophyceae,((Rhodellophyceae,(Florideophyceae,Bangiophyceae)),(Porphyridiophyceae,(Compsopogonophyceae,Stylonematophyceae)))),(((Palmophyllales,Prasinococcales),(Mamiellales,Monomastigales),(Trebouxiophyceae,(Ulvophyceae,Chlorophyceae)))),((Mesostigma,Chlorokybophyceae),(Klebsormidiophyceae,(Charophyceae,(Zygnematophyta,(Coleochaetophyceae,(Bryophyta,(Marchantiophyta,(Anthoceratophyta,(Euphyllophytina,Lycopodiophyta)))))))),((Filosa,(Paramyxida,Haplosporidia),(Phytomyxea,Vampyrellidae)),(Foraminifera,(Polycystinea,Acantharea)))),((Ciliophora,((Dinoflagellata,Perkinsozoa),(Apicomplexa,Chromerida))),(Labyrinthulomycota,(Blastocystida,Opalinata),(Oomycota,Hyphochytriomycota),((Pelagophyceae,Bacillariophyceae),(Actinophrida,((Raphidophyceae,(Xanthophyceae,Phaeophyceae)),(Chrysophyta,Eustigmatophyceae,Synchromophyceae))))))));
	poly_03	(M,(O,(E1,E2,(S,A,H1,H2))));	+	-	-	+	+	+	((((Tubulinea,Discosea),(Variosea,Cutosea,Archamoebea,(Myxogastria,Protosporangiida,Dictyostelia))),((Breviatea,Apusomonadida),((Rotosphaerida,(Microsporidia,(Blastocladiomycota,((Neocallimastiqomycota,Chytridiomycota),(Zoopagomycota,(Mucoromycota,(Basidiomycota,Ascomycota)))))),((Mesomycetozoea,Coralliochytrium),(Filasterea,(Choanoflagellata,(Homoscleromorpha,Calcarea,(Hexactinellida,Demospongiae),(Placozoa,(Ctenophora,(Cnidaria,(Deuterostomia,Proterostomia)))))))))),((Jakobida,(Heterolobosea,(Euglenida,(Diplonemida,Kinetoplastea)))),(Preaxostyla,(Parabasalia,(Caviomonadidae,(Chilomastix,(Kipferlia,(Dysnectes,(Diplomonadida,Retortamonas)))))))),((((Katablepharidea,(Cryptophyceae,Goniomonadida)),(Pavlovales,(Phaeocystales,(Prymnesiales,(Coccolithales,Isochrysidales))))),(Glaucophyta,(Cyanidiophyceae,((Rhodellophyceae,(Florideophyceae,Bangiophyceae)),(Porphyridiophyceae,(Compsopogonophyceae,Stylonematophyceae))))),(((Palmophyllales,Prasinococcales),(Mamiellales,Monomastigales),(Trebouxiophyceae,(Ulvophyceae,Chlorophyceae)))),((Mesostigma,Chlorokybophyceae),(Klebsormidiophyceae,(Charophyceae,(Zygnematophyta,(Coleochaetophyceae,(Bryophyta,(Marchantiophyta,(Anthoceratophyta,(Euphyllophytina,Lycopodiophyta)))))))),((Filosa,(Paramyxida,Haplosporidia),(Phytomyxea,Vampyrellidae)),(Foraminifera,(Polycystinea,Acantharea)))),((Ciliophora,((Dinoflagellata,Perkinsozoa),(Apicomplexa,Chromerida))),(Labyrinthulomycota,(Blastocystida,Opalinata),(Oomycota,Hyphochytriomycota),((Pelagophyceae,Bacillariophyceae),(Actinophrida,((Raphidophyceae,(Xanthophyceae,Phaeophyceae)),(Chrysophyta,Eustigmatophyceae,Synchromophyceae))))))));
EVOSEA (WITHIN AMOEBOZOA)	dicho_04	(M1,(M2,(O,(E,(S,(A,H)))))));	+	+	+	-	+	+	(((Cutosea,Archamoebea),(Variosea,(Myxogastria,(Protosporangiida,Dictyostelia)))),((Tubulinea,Discosea),((Breviatea,Apusomonadida),((Rotosphaerida,(Microsporidia,(Blastocladiomycota,((Neocallimastiqomycota,Chytridiomycota),(Zoopagomycota,(Mucoromycota,(Basidiomycota,Ascomycota)))))),((Mesomycetozoea,Coralliochytrium),(Filasterea,(Choanoflagellata,((Hexactinellida,Demospongiae),((Homoscleromorpha,Calcarea),(Placozoa,(Ctenophora,(Cnidaria,(Deuterostomia,Proterostomia)))))))))),((Jakobida,(Heterolobosea,(Euglenida,(Diplonemida,Kinetoplastea)))),(Preaxostyla,(Parabasalia,(Caviomonadidae,(Chilomastix,(Kipferlia,(Dysnectes,(Diplomonadida,Retortamonas)))))))),((((Filosa,((Paramyxida,Haplosporidia),(Phytomyxea,Vampyrellidae)),(Foraminifera,(Polycystinea,Acantharea)))),((Ciliophora,((Dinoflagellata,Perkinsozoa),(Apicomplexa,Chromerida))),(Labyrinthulomycota,((Blastocystida,Opalinata),(Oomycota,Hyphochytriomycota),((Pelagophyceae,Bacillariophyceae),(Actinophrida,((Raphidophyceae,(Xanthophyceae,Phaeophyceae)),(Chrysophyta,Eustigmatophyceae,Synchromophyceae)))))))),((((Katablepharidea,(Cryptophyceae,Goniomonadida)),(Pavlovales,(Phaeocystales,(Prymnesiales,(Coccolithales,Isochrysidales))))),(((Cyanidiophyceae,((Rhodellophyceae,(Florideophyceae,Bangiophyceae)),(Porphyridiophyceae,(Compsopogonophyceae,Stylonematophyceae)))),Glaucophyta),(((Palmophyllales,Prasinococcales),(Mamiellales,Monomastigales),(Trebouxiophyceae,(Ulvophyceae,Chlorophyceae)))),((Mesostigma,Chlorokybophyceae),(Klebsormidiophyceae,(Charophyceae,(Coleochaetophyceae,(Zygnematophyta,(Bryophyta,(Marchantiophyta,(Anthoceratophyta,(Euphyllophytina,Lycopodiophyta))))))))))));
	poly_04	(M1,(M2,M3,(O,(E1,E2(S,A,H1,H2)))));	+	-	-	-	+	+	((((Variosea,Cutosea,Archamoebea,(Myxogastria,Protosporangiida,Dictyostelia)),(Tubulinea,Discosea),((Breviatea,Apusomonadida),((Rotosphaerida,(Microsporidia,(Blastocladiomycota,((Neocallimastiqomycota,Chytridiomycota),(Zoopagomycota,(Mucoromycota,(Basidiomycota,Ascomycota)))))),((Mesomycetozoea,Coralliochytrium),(Filasterea,(Choanoflagellata,(Homoscleromorpha,Calcarea,(Hexactinellida,Demospongiae),(Placozoa,(Ctenophora,(Cnidaria,(Deuterostomia,Proterostomia)))))))))),((Jakobida,(Heterolobosea,(Euglenida,(Diplonemida,Kinetoplastea)))),(Preaxostyla,(Parabasalia,(Caviomonadidae,(Chilomastix,(Kipferlia,(Dysnectes,(Diplomonadida,Retortamonas)))))))),((((Katablepharidea,(Cryptophyceae,Goniomonadida)),(Pavlovales,(Phaeocystales,(Prymnesiales,(Coccolithales,Isochrysidales))))),(Glaucophyta,(Cyanidiophyceae,((Rhodellophyceae,(Florideophyceae,Bangiophyceae)),(Porphyridiophyceae,(Compsopogonophyceae,Stylonematophyceae)))))),(((Palmophyllales,Prasinococcales),(Mamiellales,Monomastigales),(Trebouxiophyceae,(Ulvophyceae,Chlorophyceae)))),((Mesostigma,Chlorokybophyceae),(Klebsormidiophyceae,(Charophyceae,(Zygnematophyta,(Coleochaetophyceae,(Bryophyta,(Marchantiophyta,(Anthoceratophyta,(Euphyllophytina,Lycopodiophyta)))))))),((Filosa,(Paramyxida,Haplosporidia),(Phytomyxea,Vampyrellidae)),(Foraminifera,(Polycystinea,Acantharea)))),((Ciliophora,((Dinoflagellata,Perkinsozoa),(Apicomplexa,Chromerida))),(Labyrinthulomycota,(Blastocystida,Opalinata),(Oomycota,Hyphochytriomycota),((Pelagophyceae,Bacillariophyceae),(Actinophrida,((Raphidophyceae,(Xanthophyceae,Phaeophyceae)),(Chrysophyta,Eustigmatophyceae,Synchromophyceae))))))));
TUBULINEA (WITHIN AMOEBOZOA)	dicho_05	(M1,(M2,(M3,(O,(E,(S,(A,H)))))));	+	+	+	-	+	+	(Tubulinea,(Discosea,(((Cutosea,Archamoebea),(Variosea,(Myxogastria,(Protosporangiida,Dictyostelia)))),(((Breviatea,Apusomonadida),((Rotosphaerida,(Microsporidia,(Blastocladiomycota,((Neocallimastiqomycota,Chytridiomycota),(Zoopagomycota,(Mucoromycota,(Basidiomycota,Ascomycota)))))),((Mesomycetozoea,Coralliochytrium),(Filasterea,(Choanoflagellata,((Hexactinellida,Demospongiae),((Homoscleromorpha,Calcarea),(Placozoa,(Ctenophora,(Cnidaria,(Deuterostomia,Proterostomia)))))))))),((Jakobida,(Heterolobosea,(Euglenida,(Diplonemida,Kinetoplastea)))),(Preaxostyla,(Parabasalia,(Caviomonadidae,(Chilomastix,(Kipferlia,(Dysnectes,(Diplomonadida,Retortamonas)))))))),((((Filosa,((Paramyxida,Haplosporidia),(Phytomyxea,Vampyrellidae)),(Foraminifera,(Polycystinea,Acantharea)))),((Ciliophora,((Dinoflagellata,Perkinsozoa),(Apicomplexa,Chromerida))),(Labyrinthulomycota,((Blastocystida,Opalinata),(Oomycota,Hyphochytriomycota),((Pelagophyceae,Bacillariophyceae),(Actinophrida,((Raphidophyceae,(Xanthophyceae,Phaeophyceae)),(Chrysophyta,Eustigmatophyceae,Synchromophyceae)))))))),((((Katablepharidea,(Cryptophyceae,Goniomonadida)),(Pavlovales,(Phaeocystales,(Prymnesiales,(Coccolithales,Isochrysidales))))),(((Cyanidiophyceae,((Rhodellophyceae,(Florideophyceae,Bangiophyceae)),(Porphyridiophyceae,(Compsopogonophyceae,Stylonematophyceae)))),Glaucophyta),(((Palmophyllales,Prasinococcales),(Mamiellales,Monomastigales),(Trebouxiophyceae,(Ulvophyceae,Chlorophyceae)))),((Mesostigma,Chlorokybophyceae),(Klebsormidiophyceae,(Charophyceae,(Coleochaetophyceae,(Zygnematophyta,(Bryophyta,(Marchantiophyta,(Anthoceratophyta,(Euphyllophytina,Lycopodiophyta))))))))))));
	poly_05	(M1,(M2,M3,(O,(E1,E2(S,A,H1,H2)))));	+	-	-	-	+	+	(Tubulinea,(Discosea,(Variosea,Cutosea,Archamoebea,(Myxogastria,Protosporangiida,Dictyostelia)),((Breviatea,Apusomonadida,((Rotosphaerida,(Microsporidia,(Blastocladiomycota,((Neocallimastiqomycota,Chytridiomycota),(Zoopagomycota,(Mucoromycota,(Basidiomycota,Ascomycota)))))),((Mesomycetozoea,Coralliochytrium),(Filasterea,(Choanoflagellata,(Homoscleromorpha,Calcarea,(Hexactinellida,Demospongiae),(Placozoa,(Ctenophora,(Cnidaria,(Deuterostomia,Proterostomia)))))))))),((Jakobida,(Heterolobosea,(Euglenida,(Diplonemida,Kinetoplastea)))),(Preaxostyla,(Parabasalia,(Caviomonadidae,(Chilomastix,(Kipferlia,(Dysnectes,(Diplomonadida,Retortamonas)))))))),((((Katablepharidea,(Cryptophyceae,Goniomonadida)),(Pavlovales,(Phaeocystales,(Prymnesiales,(Coccolithales,Isochrysidales))))),(Glaucophyta,(Cyanidiophyceae,((Rhodellophyceae,(Florideophyceae,Bangiophyceae)),(Porphyridiophyceae,(Compsopogonophyceae,Stylonematophyceae)))))),(((Palmophyllales,Prasinococcales),(Mamiellales,Monomastigales),(Trebouxiophyceae,(Ulvophyceae,Chlorophyceae)))),((Mesostigma,Chlorokybophyceae),(Klebsormidiophyceae,(Charophyceae,(Zygnematophyta,(Coleochaetophyceae,(Bryophyta,(Marchantiophyta,(Anthoceratophyta,(Euphyllophytina,Lycopodiophyta)))))))),((Filosa,(Paramyxida,Haplosporidia),(Phytomyxea,Vampyrellidae)),(Foraminifera,(Polycystinea,Acantharea)))),((Ciliophora,((Dinoflagellata,Perkinsozoa),(Apicomplexa,Chromerida))),(Labyrinthulomycota,(Blastocystida,Opalinata),(Oomycota,Hyphochytriomycota),((Pelagophyceae,Bacillariophyceae),(Actinophrida,((Raphidophyceae,(Xanthophyceae,Phaeophyceae)),(Chrysophyta,Eustigmatophyceae,Synchromophyceae))))))));
DISCOSEA (WITHIN AMOEBOZOA)	dicho_06	(M1,(M2,(M3,(O,(E,(S,(A,H)))))));	+	+	+	-	+	+	(Discosea,(Tubulinea,(((Cutosea,Archamoebea),(Variosea,(Myxogastria,(Protosporangiida,Dictyostelia)))),(((Breviatea,Apusomonadida),((Rotosphaerida,(Microsporidia,(Blastocladiomycota,((Neocallimastiqomycota,Chytridiomycota),(Zoopagomycota,(Mucoromycota,(Basidiomycota,Ascomycota)))))),((Mesomycetozoea,Coralliochytrium),(Filasterea,(Choanoflagellata,((Hexactinellida,Demospongiae),((Homoscleromorpha,Calcarea),(Placozoa,(Ctenophora,(Cnidaria,(Deuterostomia,Proterostomia)))))))))),((Jakobida,(Heterolobosea,(Euglenida,(Diplonemida,Kinetoplastea)))),(Preaxostyla,(Parabasalia,(Caviomonadidae,(Chilomastix,(Kipferlia,(Dysnectes,(Diplomonadida,Retortamonas)))))))),((((Filosa,((Paramyxida,Haplosporidia),(Phytomyxea,Vampyrellidae)),(Foraminifera,(Polycystinea,Acantharea)))),((Ciliophora,((Dinoflagellata,Perkinsozoa),(Apicomplexa,Chromerida))),(Labyrinthulomycota,((Blastocystida,Opalinata),(Oomycota,Hyphochytriomycota),((Pelagophyceae,Bacillariophyceae),(Actinophrida,((Raphidophyceae,(Xanthophyceae,Phaeophyceae)),(Chrysophyta,Eustigmatophyceae,Synchromophyceae)))))))),((((Katablepharidea,(Cryptophyceae,Goniomonadida)),(Pavlovales,(Phaeocystales,(Prymnesiales,(Coccolithales,Isochrysidales))))),(((Cyanidiophyceae,((Rhodellophyceae,(Florideophyceae,Bangiophyceae)),(Porphyridiophyceae,(Compsopogonophyceae,Stylonematophyceae)))),Glaucophyta),(((Palmophyllales,Prasinococcales),(Mamiellales,Monomastigales),(Trebouxiophyceae,(Ulvophyceae,Chlorophyceae)))),((Mesostigma,Chlorokybophyceae),(Klebsormidiophyceae,(Charophyceae,(Coleochaetophyceae,(Zygnematophyta,(Bryophyta,(Marchantiophyta,(Anthoceratophyta,(Euphyllophytina,Lycopodiophyta))))))))))));
	poly_06	(M1,(M2,M3,(O,(E1,E2(S,A,H1,H2)))));	+	-	-	-	+	+	(Discosea,(Tubulinea,(Variosea,Cutosea,Archamoebea,(Myxogastria,Protosporangiida,Dictyostelia)),((Breviatea,Apusomonadida,((Rotosphaerida,(Microsporidia,(Blastocladiomycota,((Neocallimastiqomycota,Chytridiomycota),(Zoopagomycota,(Mucoromycota,(Basidiomycota,Ascomycota)))))),((Mesomycetozoea,Coralliochytrium),(Filasterea,(Choanoflagellata,(Homoscleromorpha,Calcarea,(Hexactinellida,Demospongiae),(Placozoa,(Ctenophora,(Cnidaria,(Deuterostomia,Proterostomia)))))))))),((Jakobida,(Heterolobosea,(Euglenida,(Diplonemida,Kinetoplastea)))),(Preaxostyla,(Parabasalia,(Caviomonadidae,(Chilomastix,(Kipferlia,(Dysnectes,(Diplomonadida,Retortamonas)))))))),((((Katablepharidea,(Cryptophyceae,Goniomonadida)),(Pavlovales,(Phaeocystales,(Prymnesiales,(Coccolithales,Isochrysidales))))),(Glaucophyta,(Cyanidiophyceae,((Rhodellophyceae,(Florideophyceae,Bangiophyceae)),(Porphyridiophyceae,(Compsopogonophyceae,Stylonematophyceae)))))),(((Palmophyllales,Prasinococcales),(Mamiellales,Monomastigales),(Trebouxiophyceae,(Ulvophyceae,Chlorophyceae)))),((Mesostigma,Chlorokybophyceae),(Klebsormidiophyceae,(Charophyceae,(Zygnematophyta,(Coleochaetophyceae,(Bryophyta,(Marchantiophyta,(Anthoceratophyta,(Euphyllophytina,Lycopodiophyta)))))))),((Filosa,(Paramyxida,Haplosporidia),(Phytomyxea,Vampyrellidae)),(Foraminifera,(Polycystinea,Acantharea)))),((Ciliophora,((Dinoflagellata,Perkinsozoa),(Apicomplexa,Chromerida))),(Labyrinthulomycota,(Blastocystida,Opalinata),(Oomycota,Hyphochytriomycota),((Pelagophyceae,Bacillariophyceae),(Actinophrida,((Raphidophyceae,(Xanthophyceae,Phaeophyceae)),(Chrysophyta,Eustigmatophyceae,Synchromophyceae))))))));
ARCHAMOEBEA (WITHIN AMOEBOZOA: EVOSEA)	dicho_07	(M1,(M2,(M3,(M4,(O,(E,(S,(A,H))))))));	+	+	+	-	+	+	(Archamoebea,(Cutosea,((Variosea,(Myxogastria,(Protosporangiida,Dictyostelia))),(Tubulinea,Discosea),(((Breviatea,Apusomonadida),((Rotosphaerida,(Microsporidia,(Blastocladiomycota,((Neocallimastiqomycota,Chytridiomycota),(Zoopagomycota,(Mucoromycota,(Basidiomycota,Ascomycota)))))),((Mesomycetozoea,Coralliochytrium),(Filasterea,(Choanoflagellata,((Hexactinellida,Demospongiae),((Homoscleromorpha,Calcarea),(Placozoa,(Ctenophora,(Cnidaria,(Deuterostomia,Proterostomia)))))))))),((Jakobida,(Heterolobosea,(Euglenida,(Diplonemida,Kinetoplastea)))),(Preaxostyla,(Parabasalia,(Caviomonadidae,(Chilomastix,(Kipferlia,(Dysnectes,(Diplomonadida,Retortamonas)))))))),((((Katablepharidea,(Cryptophyceae,Goniomonadida)),(Pavlovales,(Phaeocystales,(Prymnesiales,(Coccolithales,Isochrysidales))))),(((Cyanidiophyceae,((Rhodellophyceae,(Florideophyceae,Bangiophyceae)),(Porphyridiophyceae,(Compsopog	

									onophyceae,Stylonematophyceae)))),(Glaucophyta,(((Palmophyllales,Prasinococcales),((Mamiellales,Monomastigales),(Trebouxiophyceae,Ulvophyceae,Chlorophyceae))))),((Mesostigma,Chlorokybophyceae),(Klebsormidiophyceae,Charophyceae,Coleochaetophyceae,Zygnematophyta,Bryophyta,(Marchantiophyta,(Anthocerotophyta,(Euphyllophytina,Lycopodiophyta))))))))))))))));
	poly_07	(M1,(M2,M3,M4,(O,(E1,E2,(S,A,H1,H2)))));	+	-	-	-	+	+	(Archamoebea,(Variosea,Cutosea,(Myxoqastria,Protosporangiida,Dictyostelia),(Tubulinea,Discosea,((Breviatea,Apusomonadida,((Rotosphaerida,Microsporidia,(Blastocladiomycota,((Neocallimastigomycota,Chytridiomycota),(Zoopagomycota,(Mucoromycota,(Basidiomycota,Ascomycota))))))),((Mesomycetozoea,Corallochytrium),(Filasterea,(Choanoflagellata,(Homoscleromorpha,Calcarea,(Hexactinellida,Demospongiae),(Placozoa,(Cnidaria,Ctenophora,(Deuterostomia,Proterostomia))))))))),((Jakobida,(Heterolobosea,(Euglenida,(Diplonemida,Kinetoplastea)))),(Preaxostyla,(Parabasalia,(Caviomonadidae,(Chilomastix,(Kipferlia,(Dysnectes,(Diplomonadida,Retortamonas))))))),((Katablepharidea,(Cryptophyceae,Goniomonadida)),(Pavlovales,(Phaeocystales,(Prymnesiales,(Coccolithales,Isochrysidales))))),(Glaucophyta,(Cyanidiophyceae,((Rhodellophyceae,(Florideophyceae,Bangiophyceae)),(Porphyridiophyceae,(Compsopogonophyceae,Stylonematophyceae))))),((Palmophyllales,Prasinococcales),((Mamiellales,Monomastigales),(Trebouxiophyceae,(Ulvophyceae,Chlorophyceae)))),(Mesostigma,Chlorokybophyceae),(Klebsormidiophyceae,(Charophyceae,(Zygnematophyta,(Coleochaetophyceae,(Zygnematophyta,(Bryophyta,(Marchantiophyta,(Anthocerotophyta,(Euphyllophytina,Lycopodiophyta))))))))))))))));
HOLOZOA (ANIMALIA S.L.)	dicho_08	(O1,(O2,(O3,O4,(M,(E,(S,(A,H)))))));	+	+	+	+	-	+	(((Mesomycetozoea,Corallochytrium),(Filasterea,(Choanoflagellata,(Hexactinellida,Demospongiae),((Homoscleromorpha,Calcarea),(Placozoa,(Ctenophora,(Cnidaria,(Deuterostomia,Proterostomia))))))))),((Rotosphaerida,(Microsporidia,(Blastocladiomycota,((Neocallimastigomycota,Chytridiomycota),(Zoopagomycota,(Mucoromycota,(Basidiomycota,Ascomycota))))))),((Breviatea,Apusomonadida),(((Tubulinea,Discosea),((Cutosea,Archamoebea),(Variosea,(Myxoqastria,(Protosporangiida,Dictyostelia)))))),((Jakobida,(Heterolobosea,(Euglenida,(Diplonemida,Kinetoplastea)))),(Preaxostyla,(Parabasalia,(Caviomonadidae,(Chilomastix,(Kipferlia,(Dysnectes,(Diplomonadida,Retortamonas)))))))),((Katablepharidea,(Cryptophyceae,Goniomonadida)),(Pavlovales,(Phaeocystales,(Prymnesiales,(Coccolithales,Isochrysidales))))),(Glaucophyta,(Cyanidiophyceae,((Rhodellophyceae,(Florideophyceae,Bangiophyceae)),(Porphyridiophyceae,(Compsopogonophyceae,Stylonematophyceae))))),((Palmophyllales,Prasinococcales),((Mamiellales,Monomastigales),(Trebouxiophyceae,(Ulvophyceae,Chlorophyceae)))),(Mesostigma,Chlorokybophyceae),(Klebsormidiophyceae,(Charophyceae,(Zygnematophyta,(Coleochaetophyceae,(Zygnematophyta,(Bryophyta,(Marchantiophyta,(Anthocerotophyta,(Euphyllophytina,Lycopodiophyta))))))))))))))));
	poly_08	(O1,(O2,(O3,O4,(M,(E1,E2,(S,A,H1,H2)))));	+	-	-	+	-	+	(((Mesomycetozoea,Corallochytrium),(Filasterea,(Choanoflagellata,(Homoscleromorpha,Calcarea,(Hexactinellida,Demospongiae),(Placozoa,(Cnidaria,Ctenophora,(Deuterostomia,Proterostomia))))))),((Rotosphaerida,(Microsporidia,(Blastocladiomycota,((Neocallimastigomycota,Chytridiomycota),(Zoopagomycota,(Mucoromycota,(Basidiomycota,Ascomycota))))))),((Breviatea,Apusomonadida,((Tubulinea,Discosea),(Variosea,Cutosea,Archamoebea),(Myxogastria,Protosporangiida,Dictyostelia)))),((Jakobida,(Heterolobosea,(Euglenida,(Diplonemida,Kinetoplastea)))),(Preaxostyla,(Parabasalia,(Caviomonadidae,(Chilomastix,(Kipferlia,(Dysnectes,(Diplomonadida,Retortamonas))))))),((Katablepharidea,(Cryptophyceae,Goniomonadida)),(Pavlovales,(Phaeocystales,(Prymnesiales,(Coccolithales,Isochrysidales))))),(Glaucophyta,(Cyanidiophyceae,((Rhodellophyceae,(Florideophyceae,Bangiophyceae)),(Porphyridiophyceae,(Compsopogonophyceae,Stylonematophyceae))))),((Palmophyllales,Prasinococcales),((Mamiellales,Monomastigales),(Trebouxiophyceae,(Ulvophyceae,Chlorophyceae)))),(Mesostigma,Chlorokybophyceae),(Klebsormidiophyceae,(Charophyceae,(Zygnematophyta,(Coleochaetophyceae,(Zygnematophyta,(Bryophyta,(Marchantiophyta,(Anthocerotophyta,(Euphyllophytina,Lycopodiophyta))))))))))))))));
NUCLETMYCEA (NUCLETMYCEA, FUNGI S.L.)	dicho_09	(O1,(O2,(O3,O4,(M,(E1,E2,(S,A,H)))));	+	+	+	+	-	+	((Rotosphaerida,(Microsporidia,(Blastocladiomycota,((Neocallimastigomycota,Chytridiomycota),(Zoopagomycota,(Mucoromycota,(Basidiomycota,Ascomycota))))))),(((Mesomycetozoea,Corallochytrium),(Filasterea,(Choanoflagellata,(Hexactinellida,Demospongiae),((Homoscleromorpha,Calcarea),(Placozoa,(Cnidaria,(Ctenophora,(Deuterostomia,Proterostomia))))))))),((Breviatea,Apusomonadida,((Tubulinea,Discosea),(Variosea,Cutosea,Archamoebea),(Myxogastria,(Protosporangiida,Dictyostelia)))),((Jakobida,(Heterolobosea,(Euglenida,(Diplonemida,Kinetoplastea)))),(Preaxostyla,(Parabasalia,(Caviomonadidae,(Chilomastix,(Kipferlia,(Dysnectes,(Diplomonadida,Retortamonas))))))),(((Katablepharidea,(Cryptophyceae,Goniomonadida)),(Pavlovales,(Phaeocystales,(Prymnesiales,(Coccolithales,Isochrysidales))))),(Glaucophyta,(Cyanidiophyceae,((Rhodellophyceae,(Florideophyceae,Bangiophyceae)),(Porphyridiophyceae,(Compsopogonophyceae,Stylonematophyceae))))),((Palmophyllales,Prasinococcales),((Mamiellales,Monomastigales),(Trebouxiophyceae,(Ulvophyceae,Chlorophyceae)))),(Mesostigma,Chlorokybophyceae),(Klebsormidiophyceae,(Charophyceae,(Zygnematophyta,(Coleochaetophyceae,(Zygnematophyta,(Bryophyta,(Marchantiophyta,(Anthocerotophyta,(Euphyllophytina,Lycopodiophyta))))))))))))))));
	poly_09	(O1,(O2,(O3,O4,(M,(E1,E2,(S,A,H1,H2)))));	+	-	-	+	-	+	((Rotosphaerida,(Microsporidia,(Blastocladiomycota,((Neocallimastigomycota,Chytridiomycota),(Zoopagomycota,(Mucoromycota,(Basidiomycota,Ascomycota))))))),(((Mesomycetozoea,Corallochytrium),(Filasterea,(Choanoflagellata,(Homoscleromorpha,Calcarea,(Hexactinellida,Demospongiae),(Placozoa,(Cnidaria,Ctenophora,(Deuterostomia,Proterostomia))))))))),((Breviatea,Apusomonadida,((Tubulinea,Discosea),(Variosea,Cutosea,Archamoebea),(Myxogastria,Protosporangiida,Dictyostelia)))),((Jakobida,(Heterolobosea,(Euglenida,(Diplonemida,Kinetoplastea)))),(Preaxostyla,(Parabasalia,(Caviomonadidae,(Chilomastix,(Kipferlia,(Dysnectes,(Diplomonadida,Retortamonas))))))),(((Katablepharidea,(Cryptophyceae,Goniomonadida)),(Pavlovales,(Phaeocystales,(Prymnesiales,(Coccolithales,Isochrysidales))))),(Glaucophyta,(Cyanidiophyceae,((Rhodellophyceae,(Florideophyceae,Bangiophyceae)),(Porphyridiophyceae,(Compsopogonophyceae,Stylonematophyceae))))),((Palmophyllales,Prasinococcales),((Mamiellales,Monomastigales),(Trebouxiophyceae,(Ulvophyceae,Chlorophyceae)))),(Mesostigma,Chlorokybophyceae),(Klebsormidiophyceae,(Charophyceae,(Zygnematophyta,(Coleochaetophyceae,(Zygnematophyta,(Bryophyta,(Marchantiophyta,(Anthocerotophyta,(Euphyllophytina,Lycopodiophyta))))))))))))))));
MICROSPORIDIA (WITHIN FUNGI)	dicho_10	(O1,(O2,(O3,(O4,(O5,(M,(E,(S,(A,H))))))));	+	+	+	+	-	+	((Microsporidia,((Blastocladiomycota,((Neocallimastigomycota,Chytridiomycota),(Zoopagomycota,(Mucoromycota,(Basidiomycota,Ascomycota))))))),(Rotosphaerida,(((Mesomycetozoea,Corallochytrium),(Filasterea,(Choanoflagellata,(Hexactinellida,Demospongiae),(Homoscleromorpha,Calcarea),(Placozoa,(Ctenophora,(Cnidaria,(Deuterostomia,Proterostomia))))))))),((Breviatea,Apusomonadida),(((Tubulinea,Discosea),(Cutosea,Archamoebea),Variosea),(Myxogastria,(Protosporangiida,Dictyostelia)))),((Jakobida,(Heterolobosea,(Euglenida,(Diplonemida,Kinetoplastea)))),(Preaxostyla,(Parabasalia,(Caviomonadidae,(Chilomastix,(Kipferlia,(Dysnectes,(Diplomonadida,Retortamonas))))))),(((Katablepharidea,(Cryptophyceae,Goniomonadida)),(Pavlovales,(Phaeocystales,(Prymnesiales,(Coccolithales,Isochrysidales))))),(Glaucophyta,(Cyanidiophyceae,((Rhodellophyceae,(Florideophyceae,Bangiophyceae)),(Porphyridiophyceae,(Compsopogonophyceae,Stylonematophyceae))))),((Palmophyllales,Prasinococcales),((Mamiellales,Monomastigales),(Trebouxiophyceae,(Ulvophyceae,Chlorophyceae)))),(Mesostigma,Chlorokybophyceae),(Klebsormidiophyceae,(Charophyceae,(Zygnematophyta,(Coleochaetophyceae,(Zygnematophyta,(Bryophyta,(Marchantiophyta,(Anthocerotophyta,(Euphyllophytina,Lycopodiophyta))))))))))))))));
	poly_10	(O1,(O2,(O3,(O4,(O5,06,(M,(E1,E2,(S,A,H1,H2))))))));	+	-	-	+	-	+	((Microsporidia,((Blastocladiomycota,((Neocallimastigomycota,Chytridiomycota),(Zoopagomycota,(Mucoromycota,(Basidiomycota,Ascomycota))))))),(Rotosphaerida,(((Mesomycetozoea,Corallochytrium),(Filasterea,(Choanoflagellata,(Homoscleromorpha,Calcarea),(Hexactinellida,Demospongiae),(Placozoa,(Cnidaria,Ctenophora,(Deuterostomia,Proterostomia))))))))),((Breviatea,Apusomonadida,((Tubulinea,Discosea),(Variosea,Cutosea,Archamoebea),(Myxogastria,Protosporangiida,Dictyostelia)))),((Jakobida,(Heterolobosea,(Euglenida,(Diplonemida,Kinetoplastea)))),(Preaxostyla,(Parabasalia,(Caviomonadidae,(Chilomastix,(Kipferlia,(Dysnectes,(Diplomonadida,Retortamonas))))))),(((Katablepharidea,(Cryptophyceae,Goniomonadida)),(Pavlovales,(Phaeocystales,(Prymnesiales,(Coccolithales,Isochrysidales))))),(Glaucophyta,(Cyanidiophyceae,((Rhodellophyceae,(Florideophyceae,Bangiophyceae)),(Porphyridiophyceae,(Compsopogonophyceae,Stylonematophyceae))))),((Palmophyllales,Prasinococcales),((Mamiellales,Monomastigales),(Trebouxiophyceae,(Ulvophyceae,Chlorophyceae)))),(Mesostigma,Chlorokybophyceae),(Klebsormidiophyceae,(Charophyceae,(Zygnematophyta,(Coleochaetophyceae,(Zygnematophyta,(Bryophyta,(Marchantiophyta,(Anthocerotophyta,(Euphyllophytina,Lycopodiophyta))))))))))))))));
ROTOSPHAERIDA (WITHIN FUNGI)	dicho_11	(O1,(O2,(O3,(O4,(M,(E,(S,(A,H))))))));	+	+	+	+	-	+	((Rotosphaerida,((Microsporidia,(Blastocladiomycota,((Neocallimastigomycota,Chytridiomycota),(Zoopagomycota,(Mucoromycota,(Basidiomycota,Ascomycota))))))),(((Mesomycetozoea,Corallochytrium),(Filasterea,(Choanoflagellata,(Hexactinellida,Demospongiae),((Homoscleromorpha,Calcarea),(Placozoa,(Ctenophora,(Cnidaria,(Deuterostomia,Proterostomia))))))))),((Breviatea,Apusomonadida,(((Tubulinea,Discosea),(Cutosea,Archamoebea),(Variosea,(Myxogastria,(Protosporangiida,Dictyostelia))))),((Jakobida,(Heterolobosea,(Euglenida,(Diplonemida,Kinetoplastea)))),(Preaxostyla,(Parabasalia,(Caviomonadidae,(Chilomastix,(Kipferlia,(Dysnectes,(Diplomonadida,Retortamonas))))))),(((Katablepharidea,(Cryptophyceae,Goniomonadida)),(Pavlovales,(Phaeocystales,(Prymnesiales,(Coccolithales,Isochrysidales))))),(Glaucophyta,(Cyanidiophyceae,((Rhodellophyceae,(Florideophyceae,Bangiophyceae)),(Porphyridiophyceae,(Compsopogonophyceae,Stylonematophyceae))))),((Palmophyllales,Prasinococcales),((Mamiellales,Monomastigales),(Trebouxiophyceae,(Ulvophyceae,Chlorophyceae)))),(Mesostigma,Chlorokybophyceae),(Klebsormidiophyceae,(Charophyceae,(Zygnematophyta,(Coleochaetophyceae,(Zygnematophyta,(Bryophyta,(Marchantiophyta,(Anthocerotophyta,(Euphyllophytina,Lycopodiophyta))))))))))))))));
	poly_11	(O1,(O2,(O3,(O4,O5,(M,(E1,E2,(S,A,H1,H2))))));	+	-	-	+	-	+	((Rotosphaerida,((Microsporidia,(Blastocladiomycota,((Neocallimastigomycota,Chytridiomycota),(Zoopagomycota,(Mucoromycota,(Basidiomycota,Ascomycota))))))),(((Mesomycetozoea,Corallochytrium),(Filasterea,(Choanoflagellata,(Homoscleromorpha,Calcarea,(Hexactinellida,Demospongiae),(Placozoa,(Cnidaria,(Ctenophora,(Deuterostomia,Proterostomia))))))))),((Breviatea,Apusomonadida,(((Tubulinea,Discosea),(Cutosea,Archamoebea),(Variosea,(Myxogastria,(Protosporangiida,Dictyostelia))))),((Jakobida,(Heterolobosea,(Euglenida,(Diplonemida,Kinetoplastea)))),(Preaxostyla,(Parabasalia,(Caviomonadidae,(Chilomastix,(Kipferlia,(Dysnectes,(Diplomonadida,Retortamonas))))))),(((Katablepharidea,(Cryptophyceae,Goniomonadida)),(Pavlovales,(Phaeocystales,(Prymnesiales,(Coccolithales,Isochrysidales))))),(Glaucophyta,(Cyanidiophyceae,((Rhodellophyceae,(Florideophyceae,Bangiophyceae)),(Porphyridiophyceae,(Compsopogonophyceae,Stylonematophyceae))))),((Palmophyllales,Prasinococcales),((Mamiellales,Monomastigales),(Trebouxiophyceae,(Ulvophyceae,Chlorophyceae)))),(Mesostigma,Chlorokybophyceae),(Klebsormidiophyceae,(Charophyceae,(Zygnematophyta,(Coleochaetophyceae,(Zygnematophyta,(Bryophyta,(Marchantiophyta,(Anthocerotophyta,(Euphyllophytina,Lycopodiophyta))))))))))))))));
BLASTOCLADIOMYCOTA (WITHIN FUNGI)	dicho_12	(O1,(O2,(O3,(O4,(O5,(O6,(M,(E,(S,(A,H)))))))));	+	+	+	+	-	+	((Blastocladiomycota,((Neocallimastigomycota,Chytridiomycota),(Zoopagomycota,(Mucoromycota,(Basidiomycota,Ascomycota))))),(Microsporidia,(Rotosphaerida,(((Mesomycetozoea,Corallochytrium),(Filasterea,(Choanoflagellata,(Hexactinellida,Demospongiae),((Homoscleromorpha,Calcarea),(Placozoa,(Ctenophora,(Cnidaria,(Deuterostomia,Proterostomia))))))))),((Breviatea,Apusomonadida),(((Tubulinea,Discosea),(Cutosea,Archamoebea),(Variosea,(Myxogastria,(Protosporangiida,Dictyostelia))))),((Jakobida,(Heterolobosea,(Euglenida,(Diplonemida,Kinetoplastea)))),(Preaxostyla,(Parabasalia,(Caviomonadidae,(Chilomastix,(Kipferlia,(Dysnectes,(Diplomonadida,Retortamonas))))))),(((Katablepharidea,(Cryptophyceae,Goniomonadida)),(Pavlovales,(Phaeocystales,(Prymnesiales,(Coccolithales,Isochrysidales))))),(Glaucophyta,(Cyanidiophyceae,((Rhodellophyceae,(Florideophyceae,Bangiophyceae)),(Porphyridiophyceae,(Compsopogonophyceae,Stylonematophyceae))))),((Palmophyllales,Prasinococcales),((Mamiellales,Monomastigales),(Trebouxiophyceae,(Ulvophyceae,Chlorophyceae)))),(Mesostigma,Chlorokybophyceae),(Klebsormidiophyceae,(Charophyceae,(Zygnematophyta,(Coleochaetophyceae,(Zygnematophyta,(Bryophyta,(Marchantiophyta,(Anthocerotophyta,(Euphyllophytina,Lycopodiophyta))))))))))))))));
	poly_12	(O1,(O2,(O3,(O4,(O5,(O6,O7,(M,(E,(S,(A,H))))))))));	+	-	-	+	-	+	((Blastocladiomycota,((Neocallimastigomycota,Chytridiomycota),(Zoopagomycota,(Mucoromycota,(Basidiomycota,Ascomycota))))),(Microsporidia,(Rotosphaerida,(((Mesomycetozoea,Corallochytrium),(Filasterea,(Choanoflagellata,(Homoscleromorpha,Calcarea,(Hexactinellida,Demospongiae),(Placozoa,(Cnidaria,(Ctenophora,(Deuterostomia,Proterostomia))))))))),((Breviatea,Apusomonadida,((Tubulinea,Discosea),(Variosea,Cutosea,Archamoebea),(Myxogastria,Protosporangiida,Dictyostelia)))),((Jakobida,(Heterolobosea,(Euglenida,(Diplonemida,Kinetoplastea)))),(Preaxostyla,(Parabasalia,(Caviomonadidae,(Chilomastix,(Kipferlia,(Dysnectes,(Diplomonadida,Retortamonas))))))),(((Katablepharidea,(Cryptophyceae,Goniomonadida)),(Pavlovales,(Phaeocystales,(Prymnesiales,(Coccolithales,Isochrysidales))))),(Glaucophyta,(Cyanidiophyceae,((Rhodellophyceae,(Florideophyceae,Bangiophyceae)),(Porphyridiophyceae,(Compsopogonophyceae,Stylonematophyceae))))),((Palmophyllales,Prasinococcales),((Mamiellales,Monomastigales),(Trebouxiophyceae,(Ulvophyceae,Chlorophyceae)))),(Mesostigma,Chlorokybophyceae),(Klebsormidiophyceae,(Charophyceae,(Zygnematophyta,(Coleochaetophyceae,(Zygnematophyta,(Bryophyta,(Marchantiophyta,(Anthocerotophyta,(Euphyllophytina,Lycopodiophyta))))))))))))))));
EXCAVATA (MONOPHYLETIC)	dicho_13	(E,((O,M),(S,(A,H))));	+	+	+	+	+	+	(((Jakobida,(Heterolobosea,(Euglenida,(Diplonemida,Kinetoplastea)))),(Preaxostyla,(Parabasalia,(Caviomonadidae,(Chilomastix,(Kipferlia,(Dysnectes,(Diplomonadida,Retortamonas))))))),(((Tubulinea,Discosea),(Cutosea,Archamoebea),(Variosea,(Myxogastria,(Protosporangiida,Dictyostelia))))),((Breviatea,Apusomonadida),((Rotosphaerida,(Microsporidia,(Blastocladiomycota,((Neocallimastigomycota,Chytridiomycota),(Zoopagomycota,(Mucoromycota,(Basidiomycota,Ascomycota))))))),(((Mesomycetozoea,Corallochytrium),(Filasterea,(Choanoflagellata,((Hexactinellida,Demospongiae),(Homoscleromorpha,Calcarea),(Placozoa,(Ctenophora,(Cnidaria,(Deuterostomia,Proterostomia))))))))),((((Katablepharidea,(Cryptophyceae,Goniomonadida)),(Pavlovales,(Phaeocystales,(Prymnesiales,(Coccolithales,Isochrysidales))))),(Glaucophyta,(Cyanidiophyceae,((Rhodellophyceae,(Florideophyceae,Bangiophyceae)),(Porphyridiophyceae,(Compsopogonophyceae,Stylonematophyceae))))),((Palmophyllales,Prasinococcales),((Mamiellales,Monomastigales),(Trebouxiophyceae,(Ulvophyceae,Chlorophyceae)))),(Mesostigma,Chlorokybophyceae),(Klebsormidiophyceae,(Charophyceae,(Coleochaetophyceae,(Zygnematophyta,(Bryophyta,(Marchantiophyta,(Anthocerotophyta,(Euphyllophytina,Lycopodiophyta))))))))))))))));
	poly_13	(E,((O,M),(S,A,H1,H2)));	+	+	-	+	+	+	(((Jakobida,(Heterolobosea,(Euglenida,(Diplonemida,Kinetoplastea)))),(Preaxostyla,(Parabasalia,(Caviomonadidae,(Chilomastix,(Kipferlia,(Dysnectes,(Diplomonadida,Retortamonas))))))),(((Tubulinea,Discosea),(Variosea,Cutosea,Archamoebea),(Myxogastria,(Protosporangiida,Dictyostelia))))),((Breviatea,Apusomonadida),((Rotosphaerida,(Microsporidia,(Blastocladiomycota,((Neocallimastigomycota,Chytridiomycota),(Zoopagomycota,(Mucoromycota,(Basidiomycota,Ascomycota))))))),(((Mesomycetozoea,Corallochytrium),(Filasterea,(Choanoflagellata,(Hexactinellida,Demospongiae),(Homoscleromorpha,Calcarea),(Placozoa,(Cnidaria,(Ctenophora,(Deuterostomia,Proterostomia))))))))),((((Katablepharidea,(Cryptophyceae,Goniomonadida)),(Pavlovales,(Phaeocystales,(Prymnesiales,(Coccolithales,Isochrysidales))))),(Glaucophyta,(Cyanidiophyceae,((Rhodellophyceae,(Florideophyceae,Bangiophyceae)),(Porphyridiophyceae,(Compsopogonophyceae,Stylonematophyceae))))),((Palmophyllales,Prasinococcales),((Mamiellales,Monomastigales),(Trebouxiophyceae,(Ulvophyceae,Chlorophyceae)))),(Mesostigma,Chlorokybophyceae),(Klebsormidiophyceae,(Charophyceae,(Coleochaetophyceae,(Zygnematophyta,(Bryophyta,(Marchantiophyta,(Anthocerotophyta,(Euphyllophytina,Lycopodiophyta))))))))))))))));
DISCOBA (WITHIN EXCAVATA)	dicho_14	(E1,(E2,((O,M),(S,(A,H)))));	+	-	+	+	+	+	(((Jakobida,(Heterolobosea,(Euglenida,(Diplonemida,Kinetoplastea)))),(Preaxostyla,(Parabasalia,(Caviomonadidae,(Chilomastix,(Kipferlia,(Dysnectes,(Diplomonadida,Retortamonas))))))),(((Tubulinea,Discosea),(Cutosea,Archamoebea),(Variosea,(Myxogastria,(Protosporangiida,Dictyostelia))))),((Breviatea,Apusomonadida),((Rotosphaerida,(Microsporidia,(Blastocladiomycota,((Neocallimastigomycota,Chytridiomycota),(Zoopagomycota,(Mucoromycota,(Basidiomycota,Ascomycota))))))),(((Mesomycetozoea,Corallochytrium),(Filasterea,(Choanoflagellata,(Hexactinellida,Demospongiae),(Homoscleromorpha,Calcarea),(Placozoa,(Ctenophora,(Cnidaria,(Deuterostomia,Proterostomia))))))))),((((Katablepharidea,(Cryptophyceae,Goniomonadida)),(Pavlovales,(Phaeocystales,(Prymnesiales,(Coccolithales,Isochrysidales))))),(Glaucophyta,(Cyanidiophyceae,((Rhodellophyceae,(Florideophyceae,Bangiophyceae)),(Porphyridiophyceae,(Compsopogonophyceae,Stylonematophyceae))))),((Palmophyllales,Prasinococcales),((Mamiellales,Monomastigales),(Trebouxiophyceae,(Ulvophyceae,Chlorophyceae)))),(Mesostigma,Chlorokybophyceae),(Klebsormidiophyceae,(Charophyceae,(Coleochaetophyceae,(Zygnematophyta,(Bryophyta,(Marchantiophyta,(Anthocerotophyta,(Euphyllophytina,Lycopodiophyta))))))))))))))));
	poly_14	(E1,(E2,((O,M),(S,A,H1,H2))));	+	-	-	+	+	+	(((Jakobida,(Heterolobosea,(Euglenida,(Diplonemida,Kinetoplastea)))),(Preaxostyla,(Parabasalia,(Caviomonadidae,(Chilomastix,(Kipferlia,(Dysnectes,(Diplomonadida,Retortamonas))))))),(((Tubulinea,Discosea),(Variosea,Cutosea,Archamoebea),(Myxogastria,Protosporangiida,Dictyostelia))))),((Breviatea,Apusomonadida,((Rotosphaerida,(Microsporidia,(Blastocladiomycota,((Neocallimastigomycota,Chytridiomycota),(Zoopagomycota,(Mucoromycota,(Basidiomycota,Ascomycota))))))),(((Mesomycetozoea,Corallochytrium),(Filasterea,(Choanoflagellata,(Homoscleromorpha,Calcarea,(Hexactinellida,Demospongiae),(Placozoa,(Cnidaria,Ctenophora,(Deuterostomia,Proterostomia))))))))),((((Katablepharidea,(Cryptophyceae,Goniomonadida)),(Pavlovales,(Phaeocystales,(Prymnesiales,(Coccolithales,Isochrysidales))))),(Glaucophyta,(Cyanidiophyceae,((Rhodellophyceae,(Florideophyceae,Bangiophyceae)),(Porphyridiophyceae,(Compsopogonophyceae,Stylonematophyceae))))),((Palmophyllales,Prasinococcales),((Mamiellales,Monomastigales),(Trebouxiophyceae,(Ulvophyceae,Chlorophyceae)))),(Mesostigma,Chlorokybophyceae),(Klebsormidiophyceae,(Charophyceae,(Zygnematophyta,(Coleochaetophyceae,(Zygnematophyta,(Bryophyta,(Marchantiophyta,(Anthocerotophyta,(Euphyllophytina,Lycopodiophyta)))))))))))))));

									ta,Marchantiophyta,(Anthoceratophyta,(Euphyllophytina,Lycopodiophyta))))))))),((Filosa,(Paramyxida,Haplosporidia),(Phytomyxea,Vampyrellidae),(Foraminifera,(Polycystinea,Acantharea))),((Ciliophora,((Dinoflagellata,Perkinsozoa),(Apicomplexa,Chromerida))),(Labyrinthulomycota,(Blastocystida,Opalinata),((Oomycota,Hyphochytriomycota),((Pelagophyceae,Bacillariophyceae),(Actinophrida,((Raphidophyceae,(Xanthophyceae,Phaeophyceae)),(Chrysophyta,Eustigmatophyceae,Synchromophyceae))))))))));
NO (UNROOTED TREE)	poly_00_max	(E1,E2,(O,M),(S,A,H1,H2));	+	-	-	+	+	+	(((Discosea,Tubulinea,(Archamoebea,Cutosea,Variosea,(Myxogastria,Protosporangiida,Dictyostelia))),(Apusomonadida,Breviatea,((Rotosphaerida,(Microsporidia,(Blastocladiomycota,((Chytridiomycota,Neocallimastigomycotas),(Zopagomycota,(Mucoromycota,(Basidiomycota,Ascomycota)))))))),((Corallochytrium,Mesomycetozoea),(Filasterea,(Choanoflagellata,((Demospongiae,Hexactinellida),Calcarea,Homoscleromorpha,(Placozoa,(Cnidaria,Ctenophora,(Deuterostomia,Proterostomia)))))))))),((Preaxostyla,(Parabasalia,(Caviomonadidae,(Chilomastix,(Kipferlia,(Dysnectes,(Diplomonadida,Retortamonas)))))),(Jakobida,(Heterolobosea,(Euglenida,(Kinetoplastea,Diplonemida)))),((Glaucophyta,(Cyanidiophyceae,((Porphyridiophyceae,(Compsopogonophyceae,Stylonematophyceae)),(Rhodellophyceae,(Florideophyceae,Banqiophyceae))))),((Mesostigma,Chlorokybophyceae,(Klebsormidiophyceae,(Charophyceae,(Coleochaetophyceae,Zygnematophyta,(Bryophyta,Marchantiophyta,(Anthoceratophyta,(Lycopodiophyta,Euphyllophytina)))))))),(Palmophyllales,Prasinococcales,((Mamiellales,Monomastigales),(Trebouxiophyceae,(Ulvophyceae,Chlorophyceae)))))),(Pavlovales,(Phaeocystales,(Prymnesiales,(Coccolithales,Isochrysidales)))),(Katablepharidea,(Cryptophyceae,Goniomonadida)),((Filosa,(Paramyxida,Haplosporidia),(Phytomyxea,Vampyrellidae),(Foraminifera,(Polycystinea,Acantharea))),((Labyrinthulomycota,(Opalinata,Blastocystida),((Oomycota,Hyphochytriomycota),((Pelagophyceae,Bacillariophyceae),(Actinophrida,((Eustigmatophyceae,Synchromophyceae,Chrysophyta),(Raphidophyceae,(Phaeophyceae,Xanthophyceae))))))),(Ciliophora,((Chromerida,Apicomplexa),(Dinoflagellata,Perkinsozoa)))))));
	poly_00_max (Hacr. Exc. Mono)	(E,(O,M),(S,A,H));	+	+	+	+	+	+	(((Discosea,Tubulinea,(Archamoebea,Cutosea,Variosea,(Myxogastria,Protosporangiida,Dictyostelia))),(Apusomonadida,Breviatea,((Rotosphaerida,(Microsporidia,(Blastocladiomycota,((Chytridiomycota,Neocallimastigomycota),(Zopagomycota,(Mucoromycota,(Basidiomycota,Ascomycota)))))))),((Corallochytrium,Mesomycetozoea),(Filasterea,(Choanoflagellata,((Demospongiae,Hexactinellida),Calcarea,Homoscleromorpha,(Placozoa,(Cnidaria,Ctenophora,(Deuterostomia,Proterostomia)))))))))),((Preaxostyla,(Parabasalia,(Caviomonadidae,(Chilomastix,(Kipferlia,(Dysnectes,(Diplomonadida,Retortamonas)))))),(Jakobida,(Heterolobosea,(Euglenida,(Kinetoplastea,Diplonemida)))),((Glaucophyta,(Cyanidiophyceae,((Porphyridiophyceae,(Compsopogonophyceae,Stylonematophyceae)),(Rhodellophyceae,(Florideophyceae,Banqiophyceae))))),((Mesostigma,Chlorokybophyceae,(Klebsormidiophyceae,(Charophyceae,(Coleochaetophyceae,Zygnematophyta,(Bryophyta,Marchantiophyta,(Anthoceratophyta,(Lycopodiophyta,Euphyllophytina)))))))),(Palmophyllales,Prasinococcales,((Mamiellales,Monomastigales),(Trebouxiophyceae,(Ulvophyceae,Chlorophyceae)))))),(Pavlovales,(Phaeocystales,(Prymnesiales,(Coccolithales,Isochrysidales)))),(Katablepharidea,(Cryptophyceae,Goniomonadida)),((Filosa,(Paramyxida,Haplosporidia),(Phytomyxea,Vampyrellidae),(Foraminifera,(Polycystinea,Acantharea))),((Labyrinthulomycota,(Opalinata,Blastocystida),((Oomycota,Hyphochytriomycota),((Pelagophyceae,Bacillariophyceae),(Actinophrida,((Eustigmatophyceae,Synchromophyceae,Chrysophyta),(Raphidophyceae,(Phaeophyceae,Xanthophyceae))))))),(Ciliophora,((Chromerida,Apicomplexa),(Dinoflagellata,Perkinsozoa)))))));

Supplementary Table 5: Character matrix for the ancestral state reconstruction of the Last Eukaryotic Common Ancestor. First column has all the taxa enlisted, while in other columns six traits are annotates – (i) multinucleated state presence (1) or absence (0), presence (1) or absence (0) of sexual reproduction, polyploids known (1) or not known (0), nuclear division closed (0) or open (2), mitochondria canonical (1), hydrogenosomes (2) or mitosomes (3), and plastid primary (1) or secondary (2).

TAXON/PHYLLUM	MULTINUCLEATE	SEXUAL	POLYPLOID (>2N)	NUCLEAR DIVISION	MITOCHONDRIA	PLASTID
DINOFLAGELLATA	1	1	0	1	1	0/2
PERKINSOZOA	1	0/1	0	1	1	0
APICOMPLEXA	1	1	0	1/2	1/3	0
CHROMERIDA	0	0/1	0	1	1	2
CILIOPHORA	1	1	1	1	1/2	0/2
XANTHOPHYCEAE	1	1	0	1	1	2
PHAEOPHYCEAE	1	1	0/1	1	1	2
RAPHIDOPHYCEAE	0	1	0	1	1	2
CHRYSOPHYTA	1	1	0	2	1	2
EUSTIGMATOPHYCEAE	0	0/1	0	?	1	2
SYNCHROMOPHYCEAE	1	1	0	1	1	2
PELAGOPHYCEAE	0	0/1	0	?	1	2
BACILLARIOPHYCEAE	0/1	1	0/1	1/2	1	2
ACTINOPHRYIDA	1	1	0	1	1	0
OOMYCOTA	1	1	1	1	1	0
HYPHOCHYTRIOMYCOTA	1	0/1	0	1	1	0
BLASTOCYSTIDA	1	0/1	0/1	1	1	0
OPALINATA	1	1	0	1	1	0
LABYRINTHULOMYCOTA	1	1/?	0	2	1	0/2
FORAMINIFERA	1	1	0/1	1	1	0
POLYCYSTINEA	1	0	0	1	1	0
ACANTHAREA	1	1	0	1	1	0
PARAMYXIDA	1	0/1	0	1	1	0
HAPLOSPORIDIA	1	0/1	0/1	1	1	0
PHYTOMYXEA	1	1	0	1	1	0
VAMPYRELLIDAE	1	0/1	0	1	1	0
FILOSA	1	1	0	1/2	1	0/2
COCCOLITHALES	0	1	0	2	1	2
ISOCHRYSIDALES	0	1	0	2	1	2
PRYMNESIALES	0	1	0	2	1	2
PHAEOCYSTALES	0/1	1	0	2	1	2
PAVLOVALES	0	?	0	1	1	2
CRYPTOPHYCEAE	0	0/1	0	2	1	2
GONIOMONADIDA	0	0/1	0	?	1	0
KATABLEPHARIDEA	0	0/1	0	?	1	?
FLORIDEOPHYCEAE	1	1	0	1	1	1
BANGIOPHYCEAE	0	1	0	1	1	1
RHODELLOPHYCEAE	0	?	0	1	1	1
COMPSOPOGONOPHYCEAE	1	1	0	1	1	1
STYLONEMATOPHYCEAE	0/1	0/1	0	1	1	1
PORPHYRIDIOPHYCEAE	0	1	0	1	1	1
CYANIDIOPHYCEAE	0/?	0	0	1	1	1
EUPHYLLOPHYTINA	1	1	0/1	2	1	1
LYCOPODIOPHYTA	1	1	0	2	1	1
ANTHOCERATOPHYTA	1	1	0	2	1	1
BRYOPHYTA	1	1	0	2	1	1
MARCHANTIOPHYTA	1	1	0	2	1	1
ZYGNEMATOPHYTA	0/1	1	0/1	2	1	1
COLEOCHAETOPHYCEAE	0/1	1	0	2	1	1
CHAROPHYCEAE	1	1	0/1	2	1	1
KLEBSORMIDIOPHYCEAE	0/1	0/1	0	2	1	1
MESOSTIGMA	0	0/1	0	2	1	1
CHLOROKYBOPHYCEAE	0	0/1	0	2	1	1
ULVOPHYCEAE	1	1	0	1	1	1
CHLOROPHYCEAE	1	1	0	1	1	1
TREBOUXIOPHYCEAE	0	0/1	0	1	1	1
MAMIELLALES	0	0/1	0	?	1	1
MONOMASTIGALES	0	0/1	0	?	1	1
PALMOPHYLLALES	0/1	0/1	0	?	1	1
PRASINOCOCCALES	0/1	0/1	0	1	1	1
GLAUCOPHYTA	1	0	0	2	1	1
JAKOBIDA	0	0/1	0	2	1	0
HETEROLOBOSEA	1	0/1	0/1	1	1	0
EUGLENIDA	0	1	0	1	1	0/2
DIPLONEMIDA	0	?	0	1	1	0
KINETOPLASTEA	0	1	0/1	1	1	0
DIPLOMONADIDA	1	1	0/1	1	2/3	0
RETORTAMONAS	0	0	0	1	?	0
DYSNECTES	0	0	0	?	2	0

KIPFERLIA	0	0	0	?	2	0
CHILOMASTIX	0	0	0	1	2	0
CAVIOMONADIDAE	0	0	0	?	2	0
PARABASALIA	1	1	0	1	2	0
PREAXOSTYLA	1	1	0/1	1/2	2	0
DEUTEROSTOMIA	1	1	0/1	2	1	0
PROTEROSTOMIA	1	1	0	1/2	1	0
CNIDARIA	1	1	0	1/2	1	0
CTENOPHORA	1	1	0	2	1	0
PLACOZOA	1	0/1	0	?	1	0
HOMOSCLEROMORPHA	0/1	1	0	1/2	1	0
CALCAREA	1	1	0	1/2	1	0
HEXACTINELLIDA	1	1	0	?	1	0
DEMOSPONGIAE	1	1	0	2	1	0
CHOANOFLAGELLATA	0/1	1	0	1/2	1	0
FILASTEREA	0/1	0/1	0	?	1	0
MESOMYCETOZOEA	1	0	0	1	1	0
CORALLOCHYTRIUM	1	0/1	0	?	?	0
BREVIATEA	1	0	0	?	2	0
APUSOMONADIDA	1	0	0	?	1	0
ROTOSPHAERIDA	1	0/1	0	1	1	0
MICROSPORIDIA	1	1	0	1	3	0
NEOCALLIMASTIGOMYCOTA	1	0/1	0	1	2	0
CHYTRIDIOMYCOTA	1	1	0	1	1	0
BLASTOCLADIOMYCOTA	1	1	0/1	1	1	0
ZOOPAGOMYCOTA	1	1	0	1	1	0
MUCOROMYCOTA	1	1	0/1	1	1	0
BASIDIOMYCOTA	1	1	0/1	1	1	0
ASCOMYCOTA	1	1	0/1	1	1	0
MYXOGASTRIA	1	1	0	1/2	1	0
PROTOSPORANGIIDA	1	0	0	1	1	0
DICTYOSTELIA	1	1	0	1	1	0
VARIOSEA	1	1	0	1	1	0
CUTOSEA	0/1	0/1	0	1	1	0
ARCHAMOEBEA	1	0/1	0/1	1	2/3	0
TUBULINEA	1	1	0/1	1	1	0
DISCOSEA	1	1	0/1	1/2	1	0

Supplementary Table 6: Ancestral character state reconstruction by maximum parsimony (ordered and unordered yielded the same result). For each topology, reconstruction of the LECA (the root) characters is presented: if it was multinucleate (1 yes, 0 no, 0/1 ambiguous), sexual (0 no, 1 yes), polyploid (0 no, 1 yes), if the nuclear division is closed (1) or (2) open), what is type of mitochondria (1 canonical, 1 hydrogenosomes, 3 mitosomes) and plastid (0 absent, 1 primary or 2 secondary).

MODEL	MULTINUCLEATE	SEXUAL	POLYPLOID	NUCLEAR DIVISION	MITOCHONDRIA TYPE	PLASTID
DICHO_01	1	1	0	1	1	0
POLY_01	0or1	1	0	1	1	0
DICHO_02	1	1	0	1	1	0
POLY_02	1	0or1	0	1	1	0
DICHO_03	1	1	0	1	1	0
POLY_03	1	1	0	1	1	0
DICHO_04	1	1	0	1	1	0
POLY_04	1	1	0	1	1	0
DICHO_05	1	1	0	1	1	0
POLY_05	1	1	0	1	1	0
DICHO_06	1	1	0	1	1	0
POLY_06	1	1	0	1	1	0
DICHO_07	1	1	0	1	1or2or3	0
POLY_07	1	1	0	1	1or2or3	0
DICHO_08	1	1	0	1	1	0
POLY_08	1	0or1	0	1	1	0
DICHO_09	1	1	0	1	1	0
POLY_09	1	0or1	0	1	1	0
DICHO_10	1	1	0	1	1or3	0
POLY_10	1	1	0	1	1or3	0
DICHO_11	1	1	0	1	1	0
POLY_11	1	0or1	0	1	1	0
DICHO_12	1	1	0	1	1	0
POLY_12	1	1	0	1	1	0
DICHO_13	1	1	0	1	1	0
POLY_13	0or1	1	0	1	1	0
DICHO_14	0or1	1	0	1	1	0
POLY_14	0or1	1	0	1	1	0
POLY_00_MAX	0or1	1	0	1	1	0
POLY_00_MAX (H.E.MONO)	0or1	1	0	1	1	0

Supplementary Table 7: Ancestral character state reconstruction by maximum likelihood. Markov k-state 1 parameter model (Mk1). Rows show topology model tested, while for each character and each state likelihood in LECA is shown.

MAXIMUM LIKELIHOOD (MK1) RECONSTRUCTION

	MULTINUCLEATE		SEXUAL		POLYPLOID		NUCLEAR DIVISION		MITOCHONDRIA			PLASTID		
MODEL	0	1	0	1	0	1	OPEN	CLOSED	CANNONICAL	HYDROGENOSOME	MITOSOME	ABSENT	PRIMARY	SECONDARY
dicho_01	0.07783247	0.92216753	0.01461138	0.98538862	0.99987582	0.00012418	0.00663841	0.99336159	0.99949137	0.00033555	0.00017308	0.99938342	0.00030902	0.00030756
poly_01	0.31328064	0.68671936	0.05642788	0.94357212	0.99984668	0.00015332	0.01377169	0.98622831	0.99938669	0.0004069	0.00020641	0.99940129	0.00029699	0.00030172
dicho_02	0.0111018	0.9888982	0.15697076	0.84302924	0.99987718	0.00012282	0.00523869	0.99476131	0.99949121	0.00033364	0.00017515	0.99951364	0.00024318	0.00024318
poly_02	0.01380643	0.98619357	0.61708791	0.38291209	0.99984848	0.00015152	0.00707314	0.99292686	0.99938323	0.00040741	0.00020936	0.99940654	0.00029673	0.00029673
dicho_03	0.07783247	0.92216753	0.01461138	0.98538862	0.99987582	0.00012418	0.00663841	0.99336159	0.99949137	0.00033555	0.00017308	0.9993842	0.00030902	0.00030678
poly_03	0.0516196	0.9483804	0.05624183	0.94375817	0.99969708	0.00030292	0.00521559	0.99478441	0.99958991	0.00020912	0.00020097	0.99940681	0.00029656	0.00029663
dicho_04	0.01175933	0.98824067	0.0090403	0.9909597	0.9975311	0.0024689	0.00363394	0.99636606	0.99949571	0.0002522	0.00025209	0.99951365	0.00024318	0.00024317
poly_04	0.01206511	0.98793489	0.0063414	0.9936586	0.99969889	0.00030111	0.00402334	0.99597666	0.9996014	0.0001993	0.0001993	0.99941702	0.00029149	0.00029149
dicho_05	0.00918224	0.99081776	0.0037815	0.9962185	0.97832042	0.02167958	0.00623104	0.99376896	0.99967095	0.00016452	0.00016453	0.99952149	0.00023926	0.00023925
poly_05	0.01193588	0.98806412	0.00559035	0.99440965	0.98769373	0.01230627	0.00402244	0.99597756	0.99960426	0.00019787	0.00019787	0.99941703	0.00029149	0.00029148
dicho_06	0.00918224	0.99081776	0.0035815	0.9964185	0.97832042	0.02167958	0.0554651	0.9445349	0.99967095	0.00016452	0.00016453	0.99952149	0.00023926	0.00023925
poly_06	0.01193588	0.98806412	0.00559035	0.99440965	0.98769373	0.01230627	0.05825764	0.94174236	0.99960426	0.00019787	0.00019787	0.99941703	0.00029149	0.00029148
dicho_07	0.01727799	0.98272201	0.10987589	0.89012411	0.98898006	0.01101994	0.00320405	0.99679595	0.49358787	0.25320607	0.25320606	0.99952149	0.00023926	0.00023925
poly_07	0.01146794	0.98853206	0.06884596	0.93115404	0.98784326	0.01215674	0.00372408	0.99627592	0.49650246	0.25174877	0.25174877	0.9994172	0.0002914	0.0002914
dicho_08	0.01202138	0.98797862	0.15765103	0.84234897	0.9998772	0.0001228	0.02945587	0.97054413	0.99966197	0.00016903	0.000169	0.99951368	0.00024316	0.00024316
poly_08	0.01703769	0.98296231	0.52352616	0.47647384	0.99984663	0.00015337	0.03396358	0.96603642	0.99958656	0.00020682	0.00020662	0.99939624	0.00030188	0.00030188
dicho_09	0.01037038	0.98962962	0.11253252	0.88746748	0.9998772	0.0001228	0.00504267	0.99495733	0.99949334	0.00017515	0.00033151	0.99951379	0.0002431	0.00024311
poly_09	0.01453955	0.98546045	0.42156493	0.57843507	0.99984663	0.00015337	0.00650508	0.99349492	0.99938034	0.00021504	0.00040462	0.99939642	0.00030179	0.00030179
dicho_10	0.00918799	0.99081201	0.00727088	0.99272912	0.99987722	0.00012278	0.00321751	0.99678249	0.49354344	0.00637332	0.50008324	0.99952137	0.00023931	0.00023932
poly_10	0.01282635	0.98717365	0.02499743	0.97500257	0.99984663	0.00015337	0.00398857	0.99601143	0.49286341	0.00703495	0.50010164	0.99940672	0.00029664	0.00029664
dicho_11	0.00921755	0.99078245	0.11078474	0.88921526	0.99987855	0.00012145	0.00330349	0.99669651	0.9994999	0.00017085	0.00032925	0.99952137	0.00023932	0.00023931
poly_11	0.01286177	0.98713823	0.30626972	0.69373028	0.99984849	0.00015151	0.00411937	0.99588063	0.99938913	0.00020924	0.00040163	0.99940672	0.00029664	0.00029664
dicho_12	0.00919199	0.99080801	0.00394749	0.99605251	0.98897873	0.01102127	0.00321331	0.99678669	0.99948769	0.000177	0.00033531	0.99952137	0.00023932	0.00023931
poly_12	0.01283731	0.98716269	0.00722234	0.99277766	0.98769	0.01231	0.00398149	0.99601851	0.99937265	0.0002175	0.00040985	0.99940672	0.00029664	0.00029664
dicho_13	0.14239318	0.85760682	0.00602424	0.99397576	0.99987581	0.00012419	0.01276787	0.98723213	0.98673516	0.01277915	0.00048569	0.99938342	0.00030902	0.00030756
poly_13	0.44271592	0.55728408	0.01388683	0.98611317	0.99984661	0.00015339	0.02526709	0.97473291	0.98534336	0.0140684	0.00058824	0.99908453	0.00032316	0.00059231
dicho_14	0.39224223	0.60775777	0.01204939	0.98795061	0.99987451	0.00012549	0.06309178	0.93690822	0.98673524	0.01277907	0.00048569	0.99951167	0.00024415	0.00024418
poly_14	0.55406193	0.44593807	0.01962295	0.98037705	0.99984481	0.00015519	0.07649291	0.92350709	0.98534338	0.01406835	0.00058827	0.99939737	0.00029899	0.00030364
poly_00_max	0.42897223	0.57102777	0.00207683	0.99792317	0.99999995	0.00000005	0.00994732	0.99005268	0.99979129	0.00020579	0.0000029	0.9999895	0.00000517	0.0000053
poly_00_max (H.E.Mono)	0.22015566	0.77984434	0.00747126	0.99252874	0.99999811	0.00000189	0.0076695	0.9923305	0.99978929	0.00020238	0.0000083	0.99967994	0.00016482	0.00015524

Supplementary Table 8: Citations for multinucleated forms

Supergroup	Taxon	Citations for multinucleated forms (see Supplementary table 2 for full references)
Amoebozoa	Archamoebea	Adl et al. 2019, Pelomyxa, Daniels & Pappas 1994
Amoebozoa	Cutosea	
Amoebozoa	Dictyostelia	Bloomfield et al. 2019, Adl et al. 2019
Amoebozoa	Discosea	Band & Machnemer 1963, Byers 1979, Maciver 2016
Amoebozoa	Myxogastria	Adl et al. 2019
Amoebozoa	Protosporangiida	Adl et al. 2019
Amoebozoa	Tubulinea	Willumsen et al. 1987
Amoebozoa	Variosea	Adl et al. 2019
Archaeplastida	Anthoceratophyta	Niklas et al. 2013
Archaeplastida	Bangiophyceae	Niklas et al. 2013
Archaeplastida	Bryophyta	Niklas et al. 2013
Archaeplastida	Charophyceae	McCourt et al. 2017, Raven et al. 1979
Archaeplastida	Chlorokybophyceae	
Archaeplastida	Chlorophyceae	Higham & Bisalputra 1970, Marchant & Pickett-Heaps 1970
Archaeplastida	Coleochaetophyceae	
Archaeplastida	Compsopogonophyceae	Nichols 1964
Archaeplastida	Cyanidiophyceae	Yoon et al. 2017
Archaeplastida	Euphyllophytina	von Aderkas et al. 2005, Niklas et al. 2013
Archaeplastida	Florideophyceae	Adl et al. 2019
Archaeplastida	Glaucophyta	Steiner 2010
Archaeplastida	Klebsormidiophyceae	
Archaeplastida	Lycopodiophyta	Williams 1933, Niklas et al. 2013
Archaeplastida	Mamiellales	
Archaeplastida	Marchantiophyta	Niklas et al. 2013
Archaeplastida	*Mesostigma*	
Archaeplastida	Monomastigales	
Archaeplastida	Palmophyllales	Ballantine & Norris 1994, Adl et al. 2019
Archaeplastida	Porphyridiophyceae	
Archaeplastida	Prasinococcales	Adl et al. 2019
Archaeplastida	Rhodellophyceae	
Archaeplastida	Stylonematophyceae	Yoon et al. 2017
Archaeplastida	Trebouxiophyceae	
Archaeplastida	Ulvophyceae	Adl et al. 2019
Archaeplastida	Zygnematophyta	Hall & McCourt 2017
Discoba	Diplonemida	
Discoba	Euglenida	
Discoba	Heterolobosea	Adl et al. 2019
Discoba	Jakobida	
Discoba	Kinetoplastea	
Hacrobia	Coccolithales	
Hacrobia	Cryptophyceae	
Hacrobia	Goniomonadida	
Hacrobia	Isochrysidales	
Hacrobia	Katablepharidea	

Hacrobia	Pavlovales	
Hacrobia	Phaeocystales	Adl et al. 2019
Hacrobia	Prymnesiales	
Metamonada	Caviomonadidae	
Metamonada	*Chilomastix*	
Metamonada	Diplomonadida	Adam 2017
Metamonada	*Dysnectes*	
Metamonada	*Kipferlia*	
Metamonada	Parabasalia	Adl et al. 2019, Dolan et al. 2000A, 2000B
Metamonada	Preaxostyla	Hampl 2017
Metamonada	*Retortamonas*	
Obazoa	Apusomonadida	Heiss et al. 2017
Obazoa	Ascomycota	Spatofora et al. 2016
Obazoa	Basidiomycota	Spatofora et al. 2016
Obazoa	Blastocladiomycota	James et al. 2014
Obazoa	Breviatea	Walker et al. 2006
Obazoa	Calcarea	Franzen 1988
Obazoa	Choanoflagellata	
Obazoa	Chytridiomycota	Powell 2017
Obazoa	Cnidaria	Rosser et al. 2019, Székely et al. 2016
Obazoa	*Corallochytrium*	Raghu-Kumar 1987
Obazoa	Cristidiscoidea	Adl et al. 2019
Obazoa	Ctenophora	Hernandes-Nicaise et al. 1984
Obazoa	Demospongiae	Saller 1988, Barthel & Detmer 1990
Obazoa	Deuterostomia	Haig 2012
Obazoa	Filasterea	Sebé-Pedrós et al. 2013
Obazoa	Hexactinellida	Adl et al. 2019
Obazoa	Homoscleromorpha	
Obazoa	Mesomycetozoea	Marshall et al. 2008, Adl et al. 2019
Obazoa	Microsporidia	Cali et al. 2017
Obazoa	Mucoromycota	Spatofora et al. 2016
Obazoa	Neocallimastigomycota	Powell 2017
Obazoa	Placozoa	Ryan & Chiodin 2015
Obazoa	Proterostomia	Flemming et al. 2000, Hartenstein 1993, Tyler & Hooge 2004
Obazoa	Zoopagomycota	Spatofora et al. 2016
SAR	Acantharea	Adl et al. 2019
SAR	Actinophryida	Adl et al. 2019
SAR	Apicomplexa	Hausmann et al. 2003
SAR	Bacillariophyceae	Mann et al. 2017
SAR	Blastocystida	Kostka 2017
SAR	Chromerida	
SAR	Chrysophyta	Röderer 1979, Trenkwalder 1975
SAR	Ciliophora	Vďačný et al. 2011, Adl et al. 2019
SAR	Dinoflagellata	Guillou et al. 2008
SAR	Eustigmatophyceae	
SAR	Filosa	Ota et al. 2011
SAR	Foraminifera	Adl et al. 2019
SAR	Haplosporidia	Azevedo & Hine 2017

SAR	Hyphochytriomycota	Cook 1930, Beakes & Thines 2017
SAR	Labyrinthulomycota	Bennett et al. 2017, Azevedo & Corral 1997
SAR	Oomycota	Adl et al. 2019
SAR	Opalinata	Adl et al. 2019, Kostka 2017
SAR	Paramyxida	Lester & Hine 2017
SAR	Pelagophyceae	
SAR	Perkinsozoa	Reñé et al. 2017
SAR	Phaeophyceae	Kawai & Henry 2017, Sanders et al. 2005
SAR	Phytomyxea	Bulman & Neuhauser 2017
SAR	Polycystinea	Anderson 1976
SAR	Raphidophyceae	
SAR	Synchromophyceae	Horn et al. 2007, Koch et al. 2010
SAR	Vampyrellidae	Adl et al. 2019
SAR	Xanthophyceae	Salmaso & Tolotti 2009, Adl et al. 2019

DISCUSSION

Eukaryogenesis is one of the most interesting phenomena in the evolution of life on Earth and thus it is important to know how to approach its research. Cladistic terminology provides a frame for defining ancestral taxa (paraphyletic) and ancestral traits (plesiomorphous) among the monophyletic groups, and apomorphic traits, respectively. The first paper within this thesis (Skejo & Franjević 2020) defines the terms monophyletic (holophyletic and paraphyletic), polyphyletic, as well as apomorphic (synapomorphic and plesiomorphic), and homoplastic, in the light of eukaryotic origin and classification. It is clear that the origin of eukaryotes, despite the many questions being answered (Archibald 2011; Gould et al. 2016; Gibson et al. 2018), still represents one of the greatest unsolved mysteries in biology (López-García & Moreira 2015; Dacks et al. 2016; Martin 2017). Many historical models tried to describe the mechanism of eukaryogenesis, and many had good assumptions. However, in the light of modern findings, most of the historical hypotheses were definitely rejected. The second paper of this thesis (Tria et al. 2021) shows that duplications represent a strong phylogenetic signal from which many conclusions can be made, but also question raised, for example for the eukaryotic root. Which group is the eukaryotic root is still problematic, and this very issue makes the explanation of the eukaryotic evolutionary scenario very problematic (Zmasek & Godzik 2011; Tria et al. 2020). If the root lies within Excavata, i.e., if the excavates are paraphyletic, it is clear that modern excavate-like morphology represents the plesimorphous one. If the root is Opisthokonta, or if the root is within this group, it is then clear that the multinucleated state represents the plesiomorphic state. The third paper (Skejo et al. 2021) tests the ancestral states (sex, mitochondria, plastid, multinucleate phase) of LECA on various topologies, i.e., with different eukaryotic roots.
Three papers that make this thesis (Skejo & Franjević 2020; Tria et al. 2021; Skejo et al. 2021, see the previous chapter/s) together represent a neat and logical story, so the discussion of this thesis is going to be short.

In the following pages it is summarized which traits LECA definitely had and what are the open questions about those traits (see **LECA's traits**); then a new eukaryogenesis model in the light of this thesis' findings is briefly presented and depicted (Figure 18., see **Multinucleated model or polykaryon hypothesis**); also ancestral eukaryotic lifecycle (see **Evolution of the eukaryotic lifecycle**) is discussed, with emphasis on ubiquitous traits, such as sex, meiosis, zygote, gametes, and haploid phase. Finally, some guidelines for future research on the early eukaryote evolution are discussed (see **Future research**).

LECA'S TRAITS

The major eukaryotic traits are presented and discussed in this chapter, in order to clearly define which are ancestral and which are not. LECA (the Last Eukaryotic Common Ancestor) originated by symbiogenesis between an archaeal host and a bacterial endosymbiont (Martin & Müller 1998; Imachi et al. 2020). The major eukaryotic traits are the **nucleus**, **mitochondria**, and **sex** (Garg & Martin 2016; Gould et al. 2016) and they are herewith discussed. It would be intuitive to think that most eukaryotic membranes have an archaeal origin because an archaeon was the host, but this is, interestingly, not the case. All the eukaryotic membranes are indeed bacterial in origin and this is definitely a clue of a rare, quite astonishing, evolutionary event. The origin of eukaryotic membranes is often explained via proteobacterial vesicular activity, which consequently, during the eukaryogenesis, completely replaced the archaeal ones (Gould et al. 2016). The archaeal lineage that was involved in the eukaryogenesis as the host is now known to be a close relative of a recently discovered *Prometheoarchaeum* (Imachi et al. 2020). The bacterial lineage from which mitochondria originated is known to belong to proteobacteria but is it not clear yet whether mitochondria belong to alphaproteobacteria or represent a group sister to alphaproteobacteria (Muñoz-Gómez et al. 2022). It is, furthermore, not completely clear how did the mitochondrion enter the archaeal cell. Was it by phagotrophy (Martijn & Ettema 2013) or by some other mechanism? Since all eukaryotes have bacterial membranes, they cannot be regarded as purely archaeal descendants (Fournier & Poole 2018), but the true descendants of a polyphyletic event in which two distinct prokaryotic lineages were involved (Skejo & Franjević 2020).

All eukaryotes have at least one **nucleus**. The nucleus is eukaryotic synapomorphy. The origin of this interesting and important organelle is also not resolved yet (Martin 2005; Martin et al. 2017). The nucleus is a region within the eukaryotic cell in which all the cellular DNA has been stored, with the exceptions of mitochondrial and plastid genomes. Most eukaryotic nuclei contain more bacterial than archaeal genes, but there are exceptions, such as excavates, in which archaeal genes dominate (Brückner & Martin 2020). The chimeric nature of the nucleus is one of the strongest pieces of evidence for the polyphyletic nature of eukaryogenesis (Martin 2005; Brückner & Martin 2020). There are, interestingly, cells, such as erythrocytes in mammalian blood, which lose their nucleus during the hematopoiesis. However, as erythrocytes do not have mitochondria and nucleus, i.e., they lack any DNA, it is questionable should they be regarded as cells or just extracellular vesicles (Smith 1987). The ancestral number of nuclei was hitherto questionable so this thesis has provided the first comprehensive analysis of this trait favouring multinucleate state (syncytium, plasmodium, polykaryon, or coenocyte) as ancestral (Skejo et al. 2021).

Mitochondria are, together with anucleus, ancestral to all the eukaryotic lineages, but curiously not all modern eukaryotes have them (Roger et al. 2017). Some eukaryotes have hydrogenosomes (reduced anaerobic mitochondria producing hydrogen), some have mitosomes (completely reduced mitochondria without any DNA), and some completely lack mitochondria (Roger et al. 2017; Brückner & Martin 2020). The loss of mitochondria is now known to be secondary. Our analyses show that LECA had mitochondria (Skejo et al. 2021); as well as that mitochondrion played an important role in early phases of eukaryogenesis, i.e., our analyses favour mitochondria-early hypothesis (Tria et al. 2021).

It is still not clear how did the **mitochondrion** enter the archaeal host cell. Some speculate it was by phagotrophy (Martijn & Ettema 2013), but taken into account that many endosymbionts enter and inhabit extant non-phagotrophic fungi and other eukaryotes, this hypothesis should be re-considered. Prokaryotes are common endosymbionts in eukaryotic cytosols, especially in coenocytic ones (Hoffman & Arnold 2010). Fungi with aseptate hyphae have many proteobacterial endosymbionts, but also bacteria belonging to other phyla (Hoffman & Arnold 2010). The cytosol of the deep-sea coenocytic Xenophyophores is also rich in proteobacteria, especially alphaproteo-bacteria (Hori et al. 2014). *Pelomyxa palustris* (Amoebozoa), a giant multinucleate (2–1000 nuclei) microaerobic amoeba without mitochondria, has its cytosol full of endosymbionts that functionally replace mitochondria (Gutiérrez et al. 2017). In ciliates, methanogens endosymbionts are sometimes lost, which is explained by Muller's ratchet, i.e., by the small original effective population size. This means that only a genetically rich population of endosymbionts can survive isolated in the eukaryotic cytosol (Moran 1996). Interestingly, some eukaryotes are even unable to reproduce (i.e., to form sporangia or spores) without specific endosymbionts. (Perdida-Martinez et al. 2007).

Mitochondrion was a prerequisite for **meiosis** and **sex**. One bacterial cell division costs five to twenty-five times less ATPs (10–20 billion) than pulling of eukaryotic chromosomes during mitosis (50-500 billion ATPs for ten chromosomes) (Neidhardt et al. 1990; Nogales 2001; Garg & Martin 2016). Most mitochondria, i.e., most eukaryotes today are aerobic, which is anticipated, as there is 21% of oxygen in the atmosphere, but two billion years ago, the concentration of oxygen in the atmosphere was much lower, only about 1–3%, so it is expected that ancestral mitochondria were not aerobic, but anaerobic or facultatively aerobic (Martin & Müller 1998; Zimorski et al. 2019; Imachi et al. 2020). Eukaryotes diversified into many forms with the increase of oxygen concentration and today those forms, the descendants of the great oxidation event (GOE) are dominant life forms on Earth (e.g., Lyons et al. 2014; Zimorski et al. 2019).

LECA was **sexual**, which means it was meiotic. Previously, it was thought that sex occurred many times independently because there are so many chromosome determination ways (X0, XY, ZW, etc.) and that LECA was asexual and polyploid (Maciver 2019), but modern data reject this hypothesis. **Meiosis** is now known to be ancestral to all the eukaryotes and those eukaryotes who do not exhibit recombination during the lifecycle are derived forms that secondarily lost meiotic genes (Speijer et al. 2015; Hofstratter et al. 2018). For example, there are many polyploid asexual Amoebozoa (Maciver 2016), but the ancestral trait of the supergroup is sexual reproduction.

The whole meiotic machinery, including the genes for plasmogamy, karyogamy, pairing, and recombination, is ancestral to all the eukaryotes, and some representatives secondarily lost it (Tekle et al. 2017; Hofstatter et al. 2018). Sex is an ancestral eukaryotic trait and our ancestral state reconstruction (Skejo et al. 2021) did not reject this hypothesis. Sexual LECA can be regarded as a proven scientific fact.

An open question remains how many sexes did LECA have (Heitman 2015). Animals and plants have a bisexual system, with male and female sexes clearly defined, but already in Amoebozoa triparental, i.e., the trisexual system is known (Bloomfield et al. 2019). Basidiomycota (Fungi) have so many mating types, that for some species it could be said that hundreds of sexes exist in nature, with very complex rules how nuclei can fuse (Casselton & Kües 2007; Heitman et al. 2013; Heitman 2015). A very complex sex system was probably also present in LECA, where gametes represented nuclei with a predisposition to fuse with other nuclei, and which will undergo karyogamy, polyploidization, recombination, and reduction. Such a system could be regarded as both unisexual and multisexual. Quality selection of the gamete nuclei is simultaneous with the quality selection of mitochondria that will accompany these nuclei (Radzvilavicius et al. 2016; Razdvilavicius 2016).

Interesting eukaryotic synapomorphy is definitely **ESPs** (**Eukaryotic Signature-Proteins**). These proteins were once regarded to be unique for eukaryotes and no homologs were known among the prokaryotes (Hartman & Fedorov 2002). It is now clear that some of those proteins originated already in prokaryotes, and are not so specific, i.e., not specific to eukaryotes, but instead inherited from their ancestors. For example, some viral myosins and actin-related proteins were recently discovered in Imitervirales, and are thought to originate from the time before eukaryogenesis, i.e., from the prokaryotic world (Da Cunha et al. 2022). Furthermore, eukaryotic actins found in Bacthyarchaea are speculated to be transferred by LGT (lateral gene transfer) from early eukaryotes, which could also be true for diverse archaeal proteins (Da Cunha et al. 2022). ESPs were also found in Asgard. As more prokaryotic diversity is being discovered, it is expected that more homologs to ESPs will be found in bacterial and/or archaeal lineages.

Various aspects from the research on the origin of eukaryotes are not covered in this discussion nor were they covered in this thesis. The origin of eukaryotes is a huge field where more and more papers are being published every year. Some scientists are trying to reconstruct various aspects of LECA's biochemistry, cytology, and physiology (e.g., Grau-Bové et al. 2015; Petit et al. 2018; Klim et al. 2018; Tromer et al. 2019), others investigate ancestral genomics and proteomics from modern genomes and proteomes (e.g., Newmann et al. 2019; Fu et al. 2019; Tria et al. 2021), while some answer questions about habitat, morphology, and ethology (e.g., Skejo et al. 2021; Jamy et al. 2021). All of these findings are interesting and important, and they describe parts of early eukaryotic evolution. However, a lot of work is still in front of us before it will be definitely answered which traits LECA really had and how did they come to be.

MULTINUCLEATED MODEL OR POLYKARYON HYPOTHESIS

The unicellular and uninucleate LECA hypothesis was never questioned. Every model which depicted eukaryogenesis, almost, as a rule, shows one archaeon in which one bacterium enters (Zimorski et al. 2014). The new hypothesis of eukaryogenesis presented in this thesis represents a combination of four separated hypotheses: hydrogen hypothesis explaining the physiology of the syntrophy between archaea and bacteria (Martin & Müller 1998); "E3 hypothesis" (Entangle-Engulf-Endogenise) explaining taxonomic groups involved in eukaryogenesis – Proteobacteria, and Asgard related to recently discovered *Prometheoarchaeum* (Imachi et al. 2020); hypothesis on the origin of eukaryotic membranes from bacterial vesicles (Gould et al. 2016); and the results of this very thesis.

Multinucleated forms are ubiquitous among eukaryotes (Skejo et al. 2021). Coenocyte syncytium, plasmodium, multinucleated cell, and polykaryon are all names for a cell with many nuclei. Such cells were neglected in evolutionary research, but this state was shown to be ancestral (Skejo et al 2021). Multicellular organisms are functionally also multinucleated because a single cell with a single nucleus taken away from a multicellular organism cannot survive. Here, the **multinucleated model** or **polykaryon hypothesis** is proposed and divided into five phases – (1) biofilm phase or facultative syntrophy, (2) obligatory syntrophy, (3) endosymbiosis phase, (4) mitonuclear bottleneck, and finally (5) LECA phase, the ancestor of all the modern eukaryotes (Figure 18).

Phase 1) Facultative syntrophy or biofilm phase. Syntrophy between the Asgard population and Proteobacteria started a long time ago, likely in a form of mat or biofilm, a prokaryotic city, and it was definitely a special one. Despite its specialty that gave rise to modern eukaryotes, this syntrophy was facultative and probably similar to the one observed in *Prometheoarchaeum syntrophicum* (Imachi et al. 2020). Lokiarchaeota, a group to which this *Prometheoarchaeum* belongs, are suspected to be ancestrally syntrophic (Imachi et al. 2020). We know only one Asgard syntrophy for now, but we also have only one Asgard cultivated until now. Diverse syntrophies are expected to be found in the future, as this group seems to be widespread in Earth's environments (Liu et al. 2021). In the syntrophy which gave rise to the eukaryotes, the Asgard host is expected to be heterotrophic, fermentative archaea that metabolized amino-acids, and exchanged metabolites with proteobacteria. The proteobacteria are expected to be facultatively aerobic bacteria (Martin & Müller 1998; Imachi et al. 2020) which produced vesicles and had two membranes. This endosymbiont was related to alphaproteobacteria (Fan et al. 2020; Muñoz-Gómez et al. 2022).

Phase 2) Obligatory syntrophy (FECA phase). Because of various evolutionary constraints, only archaea and bacteria who lived in syntrophy survived together in one moment in obligatory syntrophy. This obligatory syntrophy happened because of extensive LGT. Many genes from proteobacteria were transferred to the Asgard host, and simultaneously proteobacteria lost many genes, so it became completely dependent on the syntrophy. Eukaryotic nuclei are full of bacterial genes, evidencing massive LGT between Archaea and protomitochondrion, who had a much larger genome than mitochondria today. Because of this obligatory syntrophy in which populations of archaea and bacteria became completely 'addicted', i.e., depended on each other, and they fully co-evolved from this phase on, this phase is to be regarded as the **First Eukaryotic Common Ancestor**, the **FECA**. Most mitochondrial genes were transferred to the host at this moment, favoring the mitochondria-early hypothesis (Tria et al. 2021).

Phase 3) Endosymbiosis phase. At this moment biofilm or syntrophy became a cell. Proteobacteria are living inside of the archaeal cytosol, but it is not clear yet how they entered (see the previous chapter). Maybe it happened because of the archaeal fusion (Naor & Gophna 2013), and they survived serendipitously trapped inside. It is, however, clear that they "bombarded" archaeal host cells with many vesicles. Archaeal host survived with complete bacterial phospholipids production machinery that came through LGT from endosymbiont, simultaneously losing the genes for its own archaeal phospholipids (Gould et al 2016). Furthermore, the whole DNA became condensed in certain regions of the cytosol, surrounded by vesicles. These regions gave rise to nuclei.

Phase 4) Mitonuclear bottleneck. The ancestor of eukaryotes had nuclei with chimeric genomes and mitochondria with proteobacterial genomes in this phase. Nuclei divide by closed division (Skejo et al. 2021), because open division resulted in aberrant nuclei. Many nuclei and many mitochondria had not survived these divisions. Only the compatible ones pulled through the earliest evolutionary constraints. Because of that, today we find duplications characteristic to all the eukaryotes (Tria et al. 2021), as well as genes shared between the majority of mitochondrial lineages (Roger et al 2017). This phase can be called mitonuclear bottleneck because only certain eukaryotic nuclei and mitochondria from it gave rise to modern diversity. Mitonuclear compatibility is known to be important among modern eukaryotes, as well (Radzvilavicius et al. 2016).

Phase 5) LECA is the ancestor of modern eukaryotic diversity, i.e., of Archaeplastida, Amoebozoa, Excavata, Hacrobia, Opisthokonta, and SAR (see Introduction). LECA was, as already said heterotrophic, facultatively aerobic, sexual, haploid, mitochondriate, and multinucleate. Different eukaryotic lineages have different percentages of bacterial and archaeal genes. For example, in Excavata that have lost their mitochondria, the nucleus is full of archaeal genes. Probably with the loss of mitochondria, many excavates exhibited simultaneous loss of bacteria-derived genes. This is the reason why in many analyses this group comes as basal. It is the closest one to Archaea in the nuclear genome, but this is a derived trait, i.e., their synapomorphy. Excavata also lack specific duplications (Brückner & Martin 2020; Tria et al. 2021). LECA's gametes who lost the ability for plasmogamy and/or karyogamy probably gave rise to the first unicellular protists, such as flagellar excavates (Skejo et al. 2021).

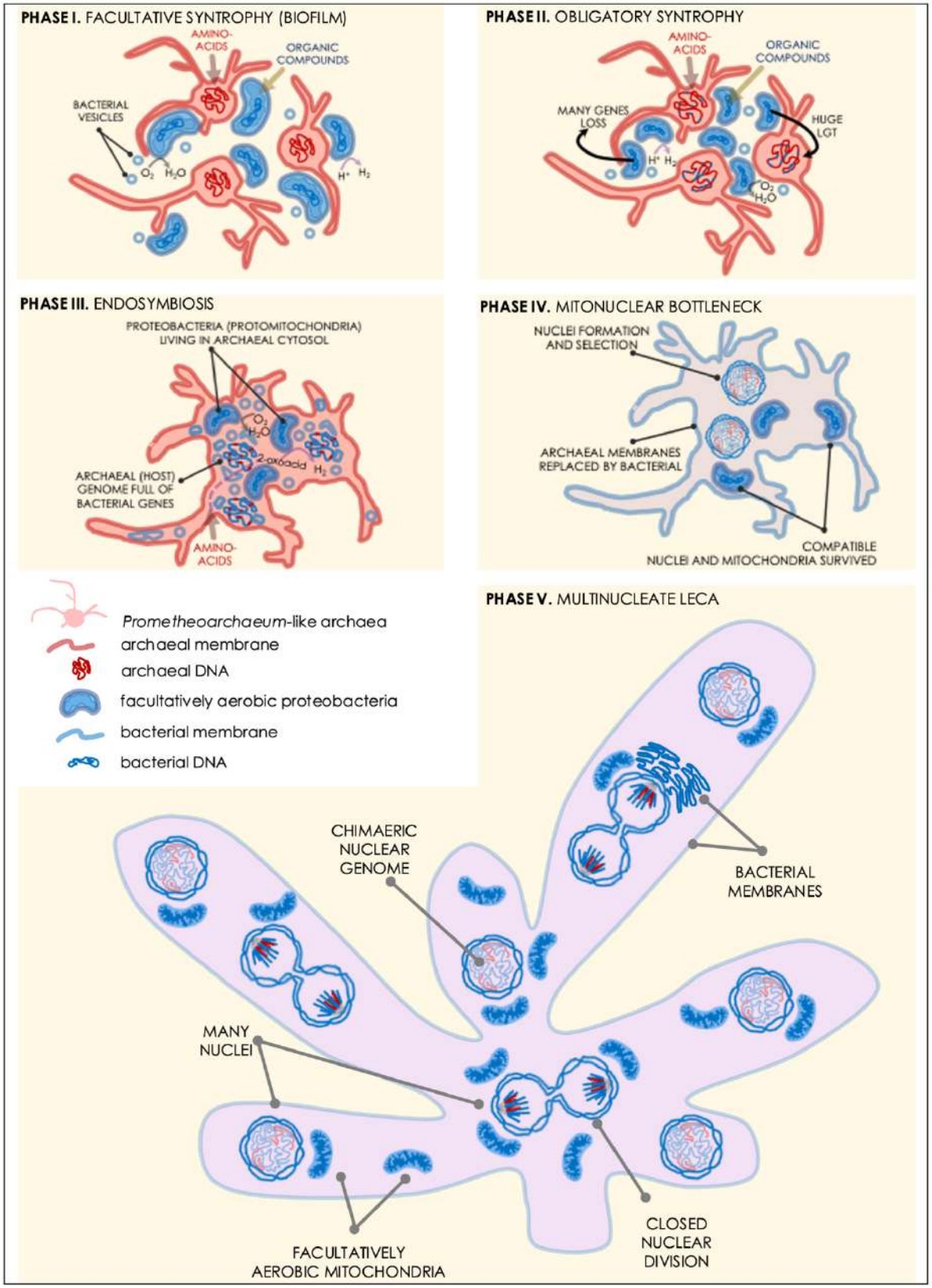


Figure 18. Schematic visualization of the polykaryon hypothesis (multinucleated model) on the eukaryotic origin. Phase I. represents biofilm of facultative syntrophy, in which proteobacteria and *Prometheoarchaeum*-like archaea exchanged metabolites. The next phase (II.) represents obligative syntrophy or the time in which proteobacteria lost many genes by functional loss, while archaea gained many genes by huge LGT. Following phase (III.) represents endosymbiosis, an event in which bacteria survived inside the archaeal cytosol. These bacteria had vesicular activity and over time, bacterial vesicles replaced all the archaeal membranes (phase IV). Finally, multinucleated, heterotrophic, sexual, and facultatively aerobic LECA; inside of which mitonuclear bottleneck happened, resulting in the survival of only compatible nuclei and mitochondria.

EVOLUTION OF THE EUKARYOTIC LIFECYCLE

Since the time of LECA, numerous modern eukaryotes have inherited parts of the ancestral life history, of course, within certain changes. Lifecycles of most of the eukaryotes include the phases (Figure 19) of the zygote, spores, gametes, and fully grown or feeding phase (Dick 1987). The feeding phase, *trophobiont,* named by Skejo et al. (2021) is not to be confused with food provider in trophobiosis, which is also named trophobiont (Delabie 2001). It is important to keep in mind that almost every organism has several 'biological forms', not only the trophobiont, but biologists most often describe only this phase because it is the easiest one to observe. From natural history angle, LECA is hence not only the syncytial phase of the first eukaryote, it is the whole ancestral eukaryotic life cycle, conserved within mitosis and meiosis present in all the eukaryotes (Garg & Martin 2016; Skejo et al. 2021).

Gametes are haploid reproductive cells, each different from another, that undergo cellular and nuclear fusion, resulting in zygote formation (Figure 19, phases 3 to 4). Spores are similar to gametes, as they are reproductive and they are haploid. The main difference is that spores usually originate by meiosis, while gametes usually originate by mitosis. As in every rule, exceptions exist (Archibald et al. 2017). Usually, one gamete harbors many mitochondria and is non-motile (e.g., egg cell), while the other usually has from one to few mitochondria and flagellar motion (e.g., spermatozoa) (Archibald et al. 2017). Zygote exhibits mitochondria quality checkpoint, so many mitochondria, for example from the egg cell, become destroyed before extensive divisions, just because of their suboptimal quality (e.g., Chappel 2013; Lieber et al. 2019). Because mitochondria mutate a lot as they have small genomes, only those with the lowest number of mutations survive with young nuclei, because all the others probably lead cells to apoptosis (Lieber et al. 2019). The zygote is ancestrally the only diploid cell in the eukaryotic lifecycle and this diploidy could be regarded as the minimal level of polyploidy, but this polyploidy is specific, because not the whole genome is duplicated, it is chimeric – two nuclei inside of one (Archibald et al. 2017). Furthermore, the zygote is ancestrally the cell in which meiosis happens. Ancestrally, meiosis comes directly after the zygote formation (e.g., Archibald et al. 2017; Vještica et al. 2021). Zygote enters a tetraploid phase (or even higher ploidy level), undergoes recombination between the homologous chromosomes, and results in four, eight, sixteen, or more daughter nuclei/cells (Archibald et al. 2017). Every nucleus originating from this cell division (which is, again, meiosis) has a different combination of alleles in comparison to any other sister nucleus. After the meiosis, there are many haploid nuclei in the shared cytoplasm and their linages proliferate by mitosis, which is basically a clonal prolongation of one nucleus' species.

Meiosis followed by serial mitoses results consequently in a multinucleated haploid trophobiont. In multinucleate organisms, it is known that different nuclei and different endosymbionts may inhabit different 'endo-microhabitats' within an organism (e.g., Anderson et al. 2015; Archibald et al. 2017; Deveau et al. 2018). This could be the reason why a single multinucleate fungus may be so euryvalent (Aleklett & Boddy 2021, Sokol et al. 2022).

In certain eukaryotes, diploid zygote did not enter meiosis and did not reduce into several haploid nuclei, but instead, it remained diploid and entered serial mitoses. This was certainly an adaptation to many evolutionary constraints (Valero et al. 1992), and because of it, many diploid lineages may be observed among the eukaryotes today (Adl et al. 2019). Multicellular diploid organisms, such as land vascular plants, and animals, are predominantly diploid and can be regarded as colonies of zygotes. In these organisms, the zygote phase is prolonged and it is not the cell that directly undergoes meiosis. Instead, the zygote undergoes many mitoses, and after each mitosis cycle certain genes are silenced, so cells become specialized. In animals, for example, only a few cells, which originate from the early phase of embryonic development and are known as primordial germ cells, stay conserved for gametogenesis (Johnson & Alberio 2015). This can be regarded as 'late' meiosis, and in animals and plants, it occurs only when the organism is sexually active, i.e., when it is adult. The prolonged diploid phase seems to be one of the important predispositions to the colonization of land, together with the evolution of stress signaling, integument, and special respiration (Ultsch 1996; de Vries et al. 2018).

Excavates are mostly unicellular and could, on the other hand, be regarded as a prolonged gamete phase. Many excavates are flagellate (Archibald et al. 2017; Adl et al. 2019) and lack genes for karyogamy (Speijer et al. 2015; Hofstatter & Lahr 2019). They often have two or four nuclei (Archibald et al. 2017). This could be a remnant from the time when they lost those genes, but two nuclei were already 'trapped' inside one shared cytosol. Figure 19. shows hypothetic ancestral eukaryotic lifecycle, with derived parts present in modern lineages, such as prolonged zygote phase in animals and plants, as well as prolonged diploid phase in rhodophytes. The early origin of two excavate groups (Discoba and Metamonada) might be related to the survival of two populations or two aberrant generations of gametes that could not finish karyogamy.

Unicellular uninucleate eukaryotes have originated many times independently. The same is true for polyploid lineages. There are ancient origins and there are recent origins. Within Fungi, for example, Microsporidia were once regarded to be plesiomorphic, ancestral, but today they are regarded as significantly reduced zygomycete-descendants (Lee et al. 2008). Unicellular yeasts, such as *Saccharomyces cervisiae*, and *Schizosaccharomyces pombe*, are now proven to be specialized descendants of hyphal ancestors (Kis et al. 2019). Who knows how many other groups with many unicellular members originated from a multinucleated or multicellular ancestor, or have such phases which stayed under the radar of science hitherto, as was the case with the multinucleated phase of *Cyanophora paradoxa* (Steiner 2010)? Unicellular eukaryotes are either zygotes (diploid unicellulars), gametes (haploid), haplospores (haploid), or diplospores (diploid) that survived independently. A lot of meticulous research is needed before there is a database useful for the reconstruction of the eukaryotic lifecycle.

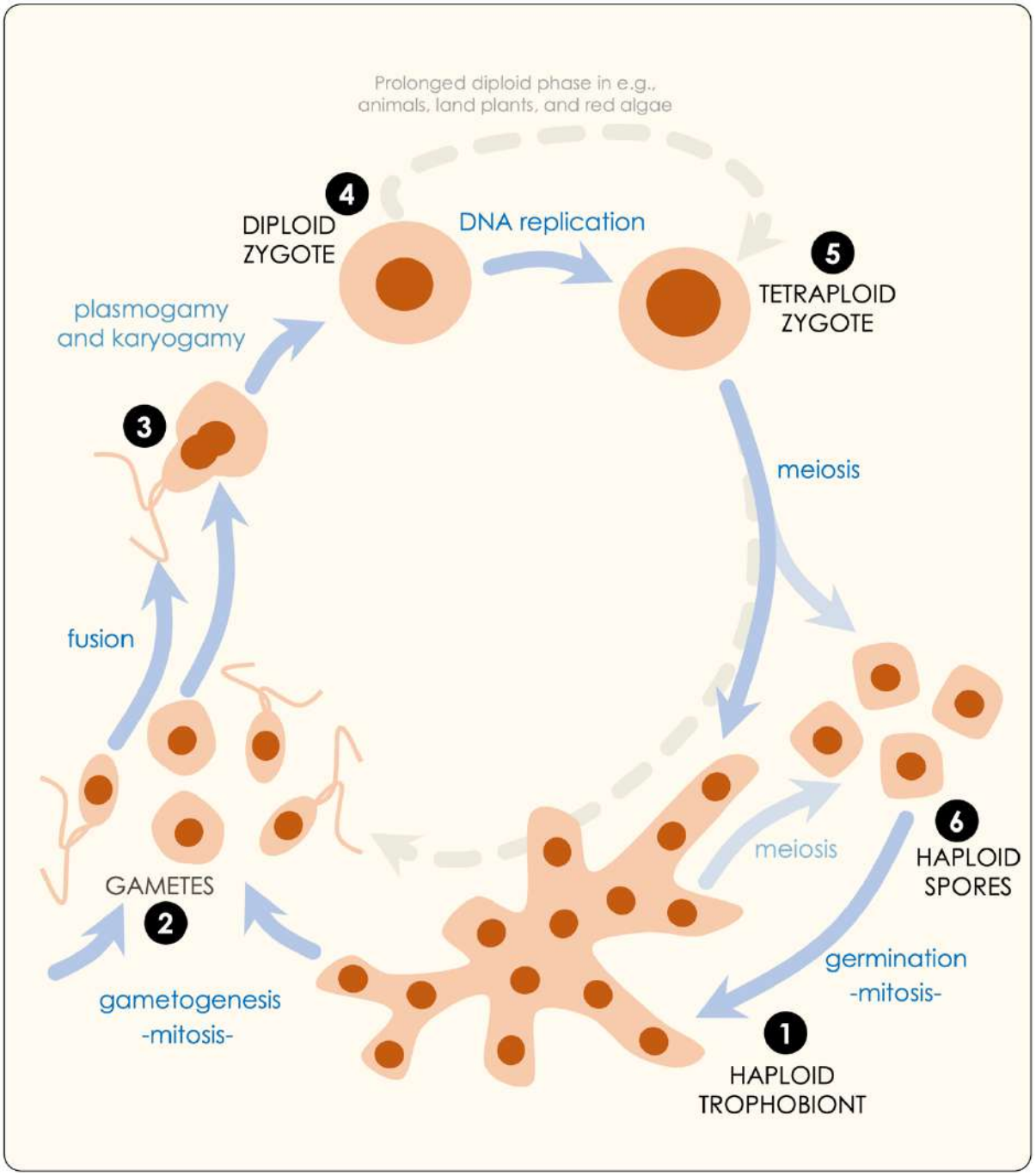


Figure 19. Schematic visualization of the hypothetic ancestral eukaryotic lifecycle. Haploid trophobiont (1) is multinucleate, heterotrophic, facultatively aerobic, mitochondriate, and sexual organisms. Certain nuclei and certain mitochondria are selected to come in packages called gametes (2). Gametes are free-living eukaryotic reproductive cells, usually dependent on water because of flagella. Some gametes have flagella (e.g., spermatozoa), while others do not (e.g., oocytes). Usually, non-motile gametes hold many mitochondria, while cytosol of motile gametes has few. Only some gametes manage to find compatible partners with whom they fuse (3) their cytoplasm (plasmogamy) and nuclei (karyogamy), forming diploid zygote (4). Zygote's nucleus undergoes one or several reduplications (5), resulting in a polyploid nucleus, which then undergoes meiosis, i.e., the division with recombination (including crossing-over) resulting in four, eight, sixteen, or more genetically different daughters. Meiosis can be the predisposition for the genetic diversity of nuclei within a multinucleated trophobiont (1), but also for the production of many different spores (6), which will consequently result in trophobionts whose nuclei are fully adapted to the environment in which the spore germinated. In animals, land plants, rhodophytes, and other eukaryotes in which the diploid phase dominates, this whole phase (i.e., the functional trophobiont) can be considered as a prolonged zygote phase. Instead of directly going into mitosis, zygotes of certain groups go through mitosis, resulting in a colony of zygotes (known as an embryo), which later has only certain cells (genetically identical to zygote) specialized for spore or gamete production. On the other hand, haploid unicellular eukaryotes could have had originated in several phases of the eukaryotic cycles (e.g., 1, 2, 3, 5), parallelly losing the others.

FUTURE RESEARCH

Deep-sea habitats hide many secrets which could contribute to the understanding of the eukaryotic origin because so many new and diverse lineages are being discovered there. However, research on these habitats is very time and money-consuming (Inagaki et al. 2015; Imachi et al. 2011, 2020). The ocean is the habitat where life originated (Martin et al. 2008), and deep-sea habitat is one of the most stable, or the least changed habitats on Earth, where many ancient prokaryotic and eukaryotic lineages survived until today. Abyssal and hadal were thought to represent the habitat where eukaryotes originated, as well, but this idea has been challenged recently, with freshwater habitats that came into the spotlight (Jamy et al. 2021). Deep-sea prokaryotes are involved in many undescribed metabolic processes (Hug et al. 2016). Metagenomics development has contributed to shedding light on parts of the undiscovered, candidate lineages, but the shadow of unknown diversity is still large and dark (Zhu et al. 2019). Cultivation of deep-sea lineages is crucial. The recent discovery of Asgard member *Prometheoarchaeum syntrophicum*, an ocean benthic prokaryote inhabiting the coastal sediment, provided a lot of data on Lokiarchaeota, as well as on the eukaryotic host (Imachi et al. 2020). However, these Lokiarchaeota are already known from a variety of habitats, including freshwater (Liu et al. 2021), hence not only do *deep-sea* lineages lack research, but all of them. Similar to the ocean prokaryotes, there are as many undiscovered deep-sea eukaryotic lineages whose existence we suspect from metagenomics (Jamy et al. 2021), but of which the scientific community still knows nothing. Xenophyophores are a great example of this not-well-understood diversity. They are huge multinucleated rhizarians which represent one of the most abundant taxonomic groups in hadal plains (Gooday et al. 2017). They have vast diversity (Buhl-Mortensen et al. 2010), but there is, at the moment of writing this thesis, no known model species, no known cultures, no available genomes sequenced. There are only about 150 to 300 eukaryotic species that are fully sequenced, most of the animals and plants (e.g., Brückner & Martin 2020; Skejo et al. 2021). Amoebozoan and fungal lineages, which are dominant in many habitats, are lacking among the model organism and sequenced genomes (Zhang et al. 2017; Gabaldón 2020). More work is needed in order to set eukaryotic model organisms for future research. Metadata for eukaryotic model organisms should be better organized and include data on morphology, physiology, and phylogeny. There should be a database with major eukaryotic traits so everybody can quickly annotate phylogenetic trees.

CONCLUSIONS

1. **Paraphyletic** means monophyletic and means ancestral. **Plesimorphous** or **plesiomorphic** means apomorphic and ancestral. Prokaryotes are monophyletic as they have a single ancestor, **LUCA** (Last Universal Common Ancestor); but they are paraphyletic towards the eukaryotes because **MITOCHONDRION** and **PLASTID** are regarded eukaryotic organelles, and not as prokaryotes. Eukaryotes are a Monophyletic and Holophyletic group because the group includes all the LECA's descendants, but they are also polyphyletic because the plastid's ancestor is a cyanobacterial lineage, which is not LECA's descendant. The conclusions have been published in Skejo & Franjević (2020).

2. Last common eukaryotic ancestor (LECA) exhibited 713 unique gene **duplications**. These duplications can resolve the topology of the eukaryotic tree. **The majority** of the duplications are **bacteria-derived** and include mitochondria-derived functions. The bacteria-derived duplications most likely originated *via* extensive **gene transfer** from the mitochondrial endosymbiont to the archaeal host's genome, so the duplications favour the mitochondria-early hypothesis. The conclusions have been published in Tria et al. (2021).

3. Eukaryotic ancestor (LECA) was **heterotrophic**, **mitochondriate**, **haploid**, **sexual,** and **multinucleated** cell (syncytial, coenocytic, plasmodial, polykaryon) with **exhibited closed nuclear division**, i.e., with the meiosis and mitosis in which nuclear membrane stays intact during the (whole) division. Uninucleate cells, previously regarded ancestral, are most likely specialized forms that originated from ancestral gametes and/or spores. Asexual, polyploid, and photosynthetic eukaryotes are also not ancestral, but derived specialized forms. The conclusions have been published in Skejo et al. (2021).

LITERATURE

A

Aanen, D. K., & Eggleton, P. (2017). Symbiogenesis: Beyond the endosymbiosis theory? Journal of Theoretical Biology, 434, 99–103.

Adl, S. et al. (2012). The revised classification of eukaryotes. Journal of Eukaryotic Microbiology, 59(5), 429–514.

Adl, S. et al. (2019). Revisions to the classification, nomenclature, and diversity of eukaryotes. Journal of Eukaryotic Microbiology, 66(1), 4–119.

Adl, S. M., et al. (2007). Diversity, nomenclature, and taxonomy of protists. Systematic Biology, 56(4), 684–689.

Albarède, F. (2009). Geochemistry: an introduction. Cambridge University Press.

Aleklett, K., & Boddy, L. (2021). Fungal behaviour: a new frontier in behavioural ecology. Trends in Ecology & Evolution, 36(9), 787–796.

Anderson, C. A. et al. (2015). Ploidy variation in multinucleate cells changes under stress. Molecular Biology of the Cell, 26(6), 1129–1140.

Angiosperm Phylogeny Group [= A.P.G.] IV. (2016). An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG IV. Botanical Journal of the Linnean Society, 181, 1–20.

Archibald, J. M. (2011). Origin of eukaryotic cells: 40 years on. Symbiosis, 54(2), 69–86.

Archibald, J. M., Simpson, A. G. B., Slamovits, C.H. (2017). Handbook of the Protists. Springer Nature.

Arndt, N. T., & Nisbet, E. G. (2012). Processes on the young Earth and the habitats of early life. Annual Review of Earth and Planetary Sciences, 40, 521–549.

B

Baross, J. A., & Hoffman, S. E. (1985). Submarine hydrothermal vents and associated gradient environments as sites for the origin and evolution of life. Origins of Life and Evolution of the Biosphere, 15(4), 327–345.

Bass, D. et al. (2009). Phylogeny of novel naked filose and reticulose Cercozoa: Granofilosea cl. n. and Proteomyxidea revised. Protist, 160, 75–109.

Baum, D. A., & Baum, B. (2014). An inside-out origin for the eukaryotic cell. BMC Biology, 12(1), 1–22.

Baum, B., & Baum, D. A. (2020). The merger that made us. BMC Biology, 18(1), 1–4.

van Bergeijk, D. A., Terlouw, B. R., Medema, M. H., & van Wezel, G. P. (2020). Ecology and genomics of Actinobacteria: new concepts for natural product discovery. Nature Reviews Microbiology, 18(10), 546–558.

Berney et al. (2015). Expansion of the 'reticulosphere': Diversity of novel branching and network-forming amoebae helps to define Variosea (Amoebozoa). Protist, 166, 271–295.

Bird, J. T. et al. (2017). Culture independent genomic comparisons reveal environmental adaptations for Altiarchaeales. Frontiers in Microbiology, 7(1221), 1–14.

Bloomfield, G. et al. (2019). Triparental inheritance in Dictyostelium. Proceedings of the National Academy of Sciences of the United States of America, 116(6), 2187–2192.

Bogorad, L. (1975). Evolution of organelles and eukaryotic genomes. Science, 188(4191), 891–898.

Booth, A., & Doolittle, W. F. (2015). Eukaryogenesis, how special really? Proceedings of the National Academy of Sciences of the United States of America, 112(33), 10278–10285.

Borges, A. R., Engstler, M., & Wolf, M. (2021). 18S rRNA gene sequence-structure phylogeny of the Trypanosomatida (Kinetoplastea, Euglenozoa) with special reference to Trypanosoma. European Journal of Protistology, 81, 125824.

Borowiec, M. L., Lee, E. K., Chiu, J. C., & Plachetzki, D. C. (2015). Extracting phylogenetic signal and accounting for bias in whole-genome data sets supports the Ctenophora as sister to remaining Metazoa. BMC Genomics, 16(1), 1–15.

Brinkmann, H., & Philippe, H. (2007). The diversity of eukaryotes and the root of the eukaryotic tree. Eukaryotic Membranes and Cytoskeleton, 20–37.

Brown, M. W. et al. (2013). Phylogenomics demonstrates that breviate flagellates are related to opisthokonts and apusomonads. Proceedings of the Royal Society B: Biological Sciences, 280(1769), 20131755.

Brückner, J., & Martin, W. F. (2020). Bacterial genes outnumber archaeal genes in eukaryotic genomes. Genome Biology and Evolution, 12(4), 282–292.

Buhl-Mortensen, L. et al. (2010). Biological structures as a source of habitat heterogeneity and biodiversity on the deep ocean margins. Marine Ecology, 31(1), 21–50.

Burki, F. et al. (2016). Untangling the early diversification of eukaryotes: A phylogenomic study of the evolutionary origins of Centrohelida, Haptophyta and Cryptista. Proceedings of the Royal Society B, 283, 20152802.

Burrows, S. M., Elbert, W., Lawrence, M. G., & Pöschl, U. (2009). Bacteria in the global atmosphere–Part 1: Review and synthesis of literature data for different ecosystems. Atmospheric Chemistry and Physics, 9(23), 9263–9280.

Butterfield, N. J. (2015). Early evolution of the Eukaryota. Palaeontology, 58(1), 5–17.

C

Casselton, L. A., & Kües, U. (2007). The origin of multiple mating types in the model mushrooms Coprinopsis cinerea and Schizophyllum commune. Sex in fungi: Molecular determination and evolutionary implications, 283–300.

Castelle, C. J. et al. (2015). Genomic expansion of domain archaea highlights roles for organisms from new phyla in anaerobic carbon cycling. Current Biology, 25(6), 690–701.

Cavalier-Smith, T. (1975). The origin of nuclei and of eukaryotic cells. Nature, 256(5517), 463–468.
Cavalier-Smith, T. (1981). Eukaryote kingdoms: seven or nine? Biosystems, 14(3–4), 461–481.
Cavalier-Smith, T. (1987). The origin of eukaryote and archaebacterial cells. Annals of the New York Academy of Sciences, 503(1), 17–54.
Cavalier-Smith, T. (1989). Archaebacteria and archezoa. Nature, 339(6220), 100–101.
Cavalier-Smith, T. (1998). A revised six-kingdom system of life. Biological Reviews, 73(3), 203–266.
Cavalier-Smith, T. (2002). The neomuran origin of archaebacteria, the negibacterial root of the universal tree and bacterial megaclassification. International Journal of Systematic and Evolutionary Microbiology, 52(1), 7–76.
Cavalier-Smith, T. (2013). Early evolution of eukaryote feeding modes, cell structural diversity, and classification of the protozoan phyla Loukozoa, Sulcozoa, and Choanozoa. European Journal of Protistology, 49(2), 115–178.
Cavalier-Smith, T. (2016). Higher classification and phylogeny of Euglenozoa. European Journal of Protistology, 56, 250–276.
Cavalier-Smith, T. et al. (2014). Multigene eukaryote phylogeny reveals the likely protozoan ancestors of opisthokonts (animals, fungi, choanozoans) and Amoebozoa. Molecular Phylogenetics and Evolution, 81, 71–85.
Cavalier-Smith, T., Chao, E. E. Y., & Oates, B. (2004). Molecular phylogeny of Amoebozoa and the evolutionary significance of the unikont *Phalansterium*. European Journal of Protistology, 40(1), 21–48.
Cavalier-Smith, T., Ema, E., & Chao, Y. (2020). Multidomain ribosomal protein trees and the planctobacterial origin of neomura (eukaryotes, archaebacteria). Protoplasma, 257(3), 621–753.
Cavalier-Smith, T. et al. (2015). Multigene phylogeny resolves deep branching of Amoebozoa. Molecular Phylogenetics and Evolution, 83, 293–304.
Cavalier-Smith, T. et al. (2016). 187-gene phylogeny of protozoan phylum Amoebozoa reveals a new class (Cutosea) of deep-branching, ultrastructurally unique, enveloped marine Lobosa and clarifies amoeba evolution. Molecular Phylogenetics and Evolution, 99, 275–296.
Cavalier-Smith, T., & Scoble, J. M. (2013). Phylogeny of Heterokonta: Incisomonas marina, a uniciliate gliding opalozoan related to *Solenicola* (Nanomonadea), and evidence that Actinophryida evolved from raphidophytes. European Journal of Protistology, 49, 328–353.
Cerón-Romero, M. A. et al. (2021) Phylogenomic analyses Of 2,786 genes in 158 lineages support a root of the eukaryotic tree of life between opisthokonts (Animals, Fungi and their microbial relatives) and all other lineages. bioRxiv. https://doi.org/10.1101/2021.02.26.433005
Chappel, S. (2013). The role of mitochondria from mature oocyte to viable blastocyst. Obstetrics and Gynecology International, 2013(183024), 1–10.
Chester, R. (2009). Marine geochemistry. John Wiley & Sons.

Clark, C. G. et al. (2006). New insights into the phylogeny of Entamoeba species provided by analysis of four new small-subunit rRNA genes. International Journal of Systematic and Evolutionary Microbiology, 56(9), 2235–2239.

CoL (2022) Catalogue of Life. Accessed 2022–02–09, available at doi:10.48580/d4tp.

Colnaghi, M., Lane, N., & Pomiankowski, A. (2020). Genome expansion in early eukaryotes drove the transition from lateral gene transfer to meiotic sex. Elife, 9, e58873.

Colp, M. J., & Archibald, J. M. (2019). Evolution: new protist predators under the sun. Current Biology, *29*(19), R936–R938.

Cordier, T., et al. (2022). Patterns of eukaryotic diversity from the surface to the deep-ocean sediment. Science Advances, 8(5), eabj9309.

Corliss, J. B. et al. (1979). Submarine thermal springs on the Galapagos rift. Science, 203, 1073–1083.

Costerton, J. W., Lewandowski, Z., Caldwell, D. E., Korber, D. R., & Lappin-Scott, H. M. (1995). Microbial biofilms. Annual Review of Microbiology, 49(1), 711–745.

Curtis, T. P., Head, I. M., Lunn, M., Woodcock, S., Schloss, P. D., & Sloan, W. T. (2006). What is the extent of prokaryotic diversity? Philosophical Transactions of the Royal Society B: Biological Sciences, 361(1475), 2023–2037.

Curtis, T. P., Sloan, W. T., & Scannell, J. W. (2002). Estimating prokaryotic diversity and its limits. Proceedings of the National Academy of Sciences of the United States of America, 99(16), 10494–10499.

D

Da Cunha, V. et al. (2022). Giant viruses encode actin-related proteins. Molecular Biology and Evolution, 39(2), msac022.

Dacks, J. B. et al. (2016). The changing view of eukaryogenesis–fossils, cells, lineages and how they all come together. Journal of Cell Science, 129(20), 3695–3703.

Deveau, A. et al. (2018). Bacterial–fungal interactions: ecology, mechanisms and challenges. FEMS Microbiology Reviews, 42(3), 335–352.

De Alda, J. A. O., Esteban, R., Diago, M. L., & Houmard, J. (2014). The plastid ancestor originated among one of the major cyanobacterial lineages. Nature Communications, 5(1), 1–10.

de Clerck, O., K. Bogaert, & Leliaert, F. (2012). Diversity and evolution of algae: primary endosymbiosis. Advances in Botanical Research, 64, 55–86.

De Duve C. 2007 The origin of eukaryotes: a reappraisal. Nature Reviews Genetics, 8, 395 – 403

de Vries, J., Curtis, B. A., Gould, S. B., & Archibald, J. M. (2018). Embryophyte stress signaling evolved in the algal progenitors of land plants. Proceedings of the National Academy of Sciences of the United States of America, 115(15), E3471–E3480.

de Vries, J., & Gould, S. B. (2018). The monoplastidic bottleneck in algae and plant evolution. Journal of Cell Science, 131(2), jcs203414.

Delabie, J. H. (2001). Trophobiosis between Formicidae and Hemiptera (Sternorrhyncha and Auchenorrhyncha): an overview. Neotropical Entomology, 30(4), 501–516.

DeLong, E. F., & Pace, N. R. (2001). Environmental diversity of bacteria and archaea. Systematic Biology, 50(4), 470–478.

Derelle, R., López-García, P., Timpano, H., & Moreira, D. (2016). A phylogenomic framework to study the diversity and evolution of stramenopiles (=heterokonts). Molecular Biology and Evolution, 33, 2890–2898.

Dick, M. W. (1987). Sexual reproduction: nuclear cycles and life-histories with particular reference to lower eukaryotes. Biological Journal of the Linnean Society, 30(2), 181–192.

Dobzhansky, T. (1973). Nothing in biology makes sense except in the light of evolution. American Biology Teacher, 3 (3), 125–129.

Dohrmann, M., & Wörheide, G. (2017). Dating early animal evolution using phylogenomic data. Scientific Reports, 7(1), 1–6.

Dombrowski, N. et al. (2020). Undinarchaeota illuminate DPANN phylogeny and the impact of gene transfer on archaeal evolution. Nature Communications, 11(1), 1–15.

E

Eikrem W. et al. (2017) Haptophyta. In: Archibald J., Simpson A., Slamovits C. (eds) Handbook of the Protists. Springer, Cham.

El-Sayed, N. M. et al. (2005). Comparative genomics of trypanosomatid parasitic protozoa. Science, 309(5733), 404–409.

Embley, T. M., & Martin, W. (2006). Eukaryotic evolution, changes and challenges. Nature, 440(7084), 623–630.

Eme, L., & Doolittle, W. F. (2015). Archaea. Current Biology, 25(19), R851–R855.

Eme, L. et al. (2017). Archaea and the origin of eukaryotes. Nature Reviews Microbiology, 15(12), 711–723.

F

Fan, L. et al. (2020). Phylogenetic analyses with systematic taxon sampling show that mitochondria branch within Alphaproteobacteria. Nature Ecology & Evolution, 4(9), 1213–1219.

Ferus, M. et al. (2017). Formation of nucleobases in a Miller–Urey reducing atmosphere. Proceedings of the National Academy of Sciences of the United States of America, 114(17), 4306–4311.

Figueroa-Martinez, F., Jackson, C., & Reyes-Prieto, A. (2019). Plastid genomes from diverse glaucophyte genera reveal a largely conserved gene content and limited architectural diversity. Genome Biology and Evolution, 11(1), 174–188.

Fournier, G. P., & Poole, A. M. (2018). A briefly argued case that Asgard archaea are part of the eukaryote tree. Frontiers in Microbiology, 9, 1896.

Fu, C. et al. (2019). Genetic and genomic evolution of sexual reproduction: echoes from LECA to the fungal kingdom. Current Opinion in Genetics & Development, 58, 70–75.

Futo, M. et al. (2021). Embryo-like features in developing *Bacillus subtilis* biofilms. Molecular Biology and Evolution, 38(1), 31–47.

G

Gabaldón, T. (2020). Grand challenges in fungal genomics and evolution. Frontiers in Fungal Biology, 1(594855), 1–3.

Garg, S. G., & Martin, W. F. (2016). Mitochondria, the cell cycle, and the origin of sex via a syncytial eukaryote common ancestor. Genome Biology and Evolution, 8(6), 1950-1970.

Garg, S. G. et al. (2021). Anomalous phylogenetic behavior of ribosomal proteins in metagenome-assembled Asgard archaea. Genome Biology and Evolution, 13(1), evaa238.

Gawryluk, R. M. et al. (2019). Non-photosynthetic predators are sister to red algae. Nature, 572(7768), 240–243.

Gibson, T. M. et al. (2018). Precise age of *Bangiomorpha pubescens* dates the origin of eukaryotic photosynthesis. Geology, 46(2), 135–138.

Gill, S., Catchpole, R., & Forterre, P. (2019). Extracellular membrane vesicles in the three domains of life and beyond. FEMS Microbiology Reviews, 43(3), 273–303.

Glenner, H. et al. (2004). Bayesian inference of the metazoan phylogeny: a combined molecular and morphological approach. Current Biology, 14(18), 1644–1649.

Gooday, A. J. et al. (2017). Giant protists (xenophyophores, Foraminifera) are exceptionally diverse in parts of the abyssal eastern Pacific licensed for polymetallic nodule exploration. Biological Conservation, 207, 106–116.

Gould, S. B., Garg, S. G., & Martin, W. F. (2016). Bacterial vesicle secretion and the evolutionary origin of the eukaryotic endomembrane system. Trends in Microbiology, 24(7), 525–534.

Gould, S. B., Waller, R. F., & McFadden, G. I. (2008). Plastid evolution. Annual Review of Plant Biology, 59, 491–517.

Grau-Bové, X., Sebé-Pedrós, A., & Ruiz-Trillo, I. (2015). The eukaryotic ancestor had a complex ubiquitin signaling system of archaeal origin. Molecular Biology and Evolution, 32(3), 726–739.

Gray, M. W. (2012). Mitochondrial evolution. Cold Spring Harbor Perspectives in Biology, 4(9), a011403.

Gray, M. W. (2014). The pre-endosymbiont hypothesis: a new perspective on the origin and evolution of mitochondria. Cold Spring Harbor Perspectives in Biology, 6(3), a016097.

Gutiérrez, G. et al. (2017). Identification of *Pelomyxa palustris* endosymbionts. Protist, 168(4), 408–424.

H

Haeckel, E. (1866). Generelle Morphologie der Organismen: allgemeine Grundzüge der organischen Formen-Wissenschaft, mechanisch begründet durch die von Charles Darwin reformirte Descendenz-Theorie. Reimer, Berlin.

Haldane, J. B. S. (1929). The origin of life. Rationalist Annual, 148, 3–10.

Hammerschmidt, K., Landan, G., Domingues Kümmel Tria, F., Alcorta, J., & Dagan, T. (2021). The order of trait emergence in the evolution of cyanobacterial multicellularity. Genome Biology and Evolution, 13(2), evaa249.

Hartman, H., & Fedorov, A. (2002). The origin of the eukaryotic cell: a genomic investigation. Proceedings of the National Academy of Sciences of the United States of America, 99(3), 1420–1425.

Hazen, R. M., & Sverjensky, D. A. (2010). Mineral surfaces, geochemical complexities, and the origins of life. Cold Spring Harbor Perspectives in Biology, 2(5), a002162.

He, D. et al. (2014). An alternative root for the eukaryote tree of life. Current Biology, 24(4), 465–470.

Heitman, J., Sun, S., & James, T. Y. (2013). Evolution of fungal sexual reproduction. Mycologia, 105(1), 1–27.

Heitman, J. (2015). Evolution of sexual reproduction: A view from the fungal kingdom supports an evolutionary epoch with sex before sexes. Fungal Biology Reviews, 29(3–4), 108–117.

Henze, K., & Martin, W. (2003). Essence of mitochondria. Nature, 426(6963), 127–128.

Hibbett, D. S. et al. (2007). A higher-level phylogenetic classification of the Fungi. Mycological Research, 111(5), 509–547.

Hoef-Emden, K., & Archibald, J. M. (2016). Cryptophyta (cryptomonads). Handbook of the Protists. Springer International Publishing, Cham, Switzerland.

Hoffman, M. T., & Arnold, A. E. (2010). Diverse bacteria inhabit living hyphae of phylogenetically diverse fungal endophytes. Applied and Environmental Microbiology, 76(12), 4063–4075.

Hofstatter, P. G., Brown, M. W., & Lahr, D. J. (2018). Comparative genomics supports sex and meiosis in diverse Amoebozoa. Genome Biology and Evolution, 10(11), 3118–3128.

Hofstatter, P. G., & Lahr, D. J. (2019). All eukaryotes are sexual, unless proven otherwise: many so-called asexuals present meiotic machinery and might be able to have sex. Bioessays, 41(6), 1800246.

Hogg, J. (1860). On the distinctions of a plant and an animal and on a fourth kingdom of Nature. The Edinburgh New Philosophical Journal (New Series), 12, 216–225.

Hori, S. et al. (2013). Active bacterial flora surrounding foraminifera (Xenophyophorea) living on the deep-sea floor. Bioscience, Biotechnology, and Biochemistry, 77(2), 381–384.

Hug, L. A. et al. (2016). A new view of the tree of life. Nature Microbiology, 1(5), 1–6.

I

Imachi, H. et al. (2011). Cultivation of methanogenic community from subseafloor sediments using a continuous–flow bioreactor. The ISME Journal, 5(12), 1913–1925.

Imachi, H. et al. (2020). Isolation of an archaeon at the prokaryote–eukaryote interface. Nature, 577(7791), 519–525.

Inagaki, F. et al. (2015). Exploring deep microbial life in coal-bearing sediment down to~ 2.5 km below the ocean floor. Science, 349(6246), 420–424.

J

Jamy, M. et al. (2021). Global patterns and rates of habitat transitions across the eukaryotic tree of life. bioRxiv, doi:10.1101/2021.11.01.466765.

Janouškovec et al. (2015). Factors mediating plastid dependency and the origins of parasitism in apicomplexans and their close relatives. Proceedings of the National Academy of Sciences of the United States of America, 112, 10200–10207.

Johnson, A. D., & Alberio, R. (2015). Primordial germ cells: the first cell lineage or the last cells standing? Development, 142(16), 2730–2739.

K

Kaiser, D. (2003). Coupling cell movement to multicellular development in myxobacteria. Nature Reviews Microbiology, 1(1), 45–54.

Kamikawa et al. (2014). Gene content evolution in discobid mitochondria deduced from the phylogenetic position and complete mitochondrial genome of *Tsukubamonas globosa*. Genome Biology and Evolution, 6, 306–315.

Katz, L. A., & Grant, J. R. (2015). Taxon-rich phylogenomic analyses resolve the eukaryotic tree of life and reveal the power of subsampling by sites. Systematic Biology, 64(3), 406–415.

Kelley, D. S. et al. (2001). An off-axis hydrothermal vent field near the Mid-Atlantic Ridge at 30 N. Nature, 412(6843), 145–149.

Kent, W. S. (1880). A manual of the Infusoria: including a description of all known Flagellate, Ciliate, and Tentaculiferous Protozoa, British and foreign, and an account of the organization and the affinities of the sponges (Vol. 1). David Bogue.

Kiss et al. (2019). Comparative genomics reveals the origin of fungal hyphae and multicellularity. Nature Communications, 10(1), 1–13.

Klim, J. et al. (2018). Ancestral state reconstruction of the apoptosis machinery in the common ancestor of eukaryotes. G3: Genes, Genomes, Genetics, 8(6), 2121–2134.

Kloesges, T. et al. (2011). Networks of gene sharing among 329 proteobacterial genomes reveal differences in lateral gene transfer frequency at different phylogenetic depths. Molecular Biology and Evolution, 28(2), 1057–1074.

Knopp, M., Stockhorst, S., van der Giezen, M., Garg, S. G., & Gould, S. B. (2021). The asgard archaeal-unique contribution to protein families of the eukaryotic common ancestor was 0.3%. Genome Biology and Evolution, 13(6), evab085.

Kondrashov, A. S. (1994). The asexual ploidy cycle and the origin of sex. Nature, 370, 213–216.

Koonin, E. V., & Martin, W. (2005). On the origin of genomes and cells within inorganic compartments. Trends in Genetics, 21(12), 647–654.

Kowallik, K. V., & Martin, W. F. (2021). The origin of symbiogenesis: An annotated English translation of Mereschkowsky's 1910 paper on the theory of two plasma lineages. Biosystems, 199, 104281.

Krabberød et al. (2017). Single cell transcriptomics, mega-phylogeny and the genetic basis of morphological innovations in Rhizaria. Molecular Biology and Evolution, 34(7), 1557–1573.

Ku, C. et al. (2015a). Endosymbiotic origin and differential loss of eukaryotic genes. Nature, 524(7566), 427–432.
Ku, C. et al. (2015b). Endosymbiotic gene transfer from prokaryotic pangenomes: Inherited chimerism in eukaryotes. Proceedings of the National Academy of Sciences of the United States of America, 112(33), 10139–10146.
Ku, C., & Martin, W. F. (2016). A natural barrier to lateral gene transfer from prokaryotes to eukaryotes revealed from genomes: the 70% rule. BMC Biology, 14(1), 1–12.
Kudryavtsev, A., & Pawlowski, J. (2013). *Squamamoeba japonica* ngn sp.(Amoebozoa): a deep-sea amoeba from the Sea of Japan with a novel cell coat structure. Protist, 164(1), 13–23.
Kumar, S., & Kumari, R. (2015). Origin, structure and function of millions of chromosomes present in the macronucleus of unicellular eukaryotic ciliate, *Oxytricha trifallax*: a model organism for transgenerationally programmed genome rearrangements. Journal of Genetics, 94(2), 171–176.

L

Lane, N., & Martin, W. F. (2015). Eukaryotes really are special, and mitochondria are why. Proceedings of the National Academy of Sciences of the United States of America, 112(35), E4823–E4823.
Laumer et al. (2019). Revisiting metazoan phylogeny with genomic sampling of all phyla. Proceedings of the Royal Society B, 286(1906), 20190831.
Lax, G. et al. (2018). Hemimastigophora is a novel supra-kingdom-level lineage of eukaryotes. Nature, 564(7736), 410–414.
Lee, S. C. et al. (2008). Microsporidia evolved from ancestral sexual fungi. Current Biology, 18(21), 1675–1679.
Leander, B. S., Lax, G., Karnkowska, A., & Simpson, A. G. B. (2017) Euglenida. In: Archibald J., Simpson A., Slamovits C. (eds) Handbook of the Protists. Springer, Cham.
Leger, M. M. et al. (2017). Organelles that illuminate the origins of Trichomonas hydrogenosomes and Giardia mitosomes. Nature Ecology and Evolution, 1, 0092.
Leliaert, F. et al. (2012). Phylogeny and molecular evolution of the green algae. Critical Reviews in Plant Sciences, 31, 1–46.
Lengeler, J. W., Drews, G., & Schlegel, H. G. (Eds.). (1999). Biology of the Prokaryotes. Georg Thieme Verlag.
Lhee, D. et al. (2019). Evolutionary dynamics of the chromatophore genome in three photosynthetic Paulinella species. Scientific Reports, 9(1), 1–11.
Li, Y. et al. (2021). A genome-scale phylogeny of the kingdom Fungi. Current Biology, 31(8), 1653–1665.
Lieber, T. et al. (2019). Mitochondrial fragmentation drives selective removal of deleterious mtDNA in the germline. Nature, 570(7761), 380–384.
Linné, C. (1735). Systemae Naturae, sive regna tria naturae, systematics proposita per classes, ordines, genera & species. Lugduni Batavorum, Haak, Leiden.

Linné, C. (1751). Philosophia Botanica: in qua explicantur fundamenta Botanica cum definitionibus partium, exemplis terminorum, observationibus rariorum, adiectis figuris aeneis. Ioannis Thomae Trattner, Caef. Reg; p. 241.

Liu, Y. et al. (2021). Expanded diversity of Asgard archaea and their relationships with eukaryotes. Nature, 593(7860), 553–557.

Loomis, W. (Ed.). (2012). The development of *Dictyostelium discoideum*. Elsevier.

López, D., Vlamakis, H., & Kolter, R. (2010). Biofilms. Cold Spring Harbor Perspectives in Biology, 2(7), a000398.

López-García, P., & Moreira, D. (2006). Selective forces for the origin of the eukaryotic nucleus. Bioessays, 28(5), 525–533.

López-García, P., & Moreira, D. (2015). Open questions on the origin of eukaryotes. Trends in Ecology & Evolution, 30(11), 697–708.

Lyons, T. W., Reinhard, C. T., & Planavsky, N. J. (2014). The rise of oxygen in Earth's early ocean and atmosphere. Nature, 506(7488), 307–315.

M

Maciver, S. K. (2016). Asexual amoebae escape Muller's ratchet through polyploidy. Trends in Parasitology, 32(11), 855–862.

Maciver, S. K. (2019). Ancestral eukaryotes reproduced asexually, facilitated by polyploidy: a hypothesis. Bioessays, 41(12), 1900152.

MacLeod, F., Kindler, G. S., Wong, H. L., Chen, R., & Burns, B. P. (2019). Asgard archaea: Diversity, function, and evolutionary implications in a range of microbiomes. AIMS Microbiology, 5(1), 48.

Makarova, K. S., Sorokin, A. V., Novichkov, P. S., Wolf, Y. I., & Koonin, E. V. (2007). Clusters of orthologous genes for 41 archaeal genomes and implications for evolutionary genomics of archaea. Biology Direct, 2(1), 1–20.

Margulis, L., Chapman, M., Guerrero, R., & Hall, J. (2006). The last eukaryotic common ancestor (LECA): acquisition of cytoskeletal motility from aerotolerant spirochetes in the Proterozoic Eon. Proceedings of the National Academy of Sciences of the United States of America, 103(35), 13080–13085.

Margulis, L., Dolan, M. F., & Guerrero, R. (2000). The chimeric eukaryote: origin of the nucleus from the karyomastigont in amitochondriate protists. Proceedings of the National Academy of Sciences of the United States of America, 97(13), 6954–6959.

Martijn, J., & Ettema, T. J. (2013). From archaeon to eukaryote: the evolutionary dark ages of the eukaryotic cell. Biochemical Society Transactions, 41(1), 451–457.

Martin, W. F. (1999). A briefly argued case that mitochondria and plastids are descendants of endosymbionts, but that the nuclear compartment is not. Proceedings of the Royal Society, London B 266, 1387 – 1395

Martin, W. (2005). Archaebacteria (Archaea) and the origin of the eukaryotic nucleus. Current Opinion in Microbiology, 8(6), 630–637.

Martin, W. F. (2017). Symbiogenesis, gradualism, and mitochondrial energy in eukaryote origin. Periodicum Biologorum, 119(3), 141–158.

Martin, W. F., Baross, J., Kelley, D., & Russell, M. J. (2008). Hydrothermal vents and the origin of life. Nature Reviews Microbiology, 6(11), 805–814.

Martin, W. F., Garg, S., & Zimorski, V. (2015). Endosymbiotic theories for eukaryote origin. Philosophical Transactions of the Royal Society B: Biological Sciences, 370(1678), 20140330.

Martin, W. F., Hoffmeister, M., Rotte, C., & Henze, K. (2001). An overview of endosymbiotic models for the origins of eukaryotes, their ATP-producing organelles (mitochondria and hydrogenosomes), and their heterotrophic lifestyle. Biological Chemistry, 382, 1521–1539.

Martin, W. F., & Müller, M. (1998). The hydrogen hypothesis for the first eukaryote. Nature, 392(6671), 37–41.

Martin, W. F., & Müller, M. (Eds.). (2007). Origin of mitochondria and hydrogenosomes. New York, NY, USA, Springer.

Martin, W. F., & Russell, M. J. (2003). On the origins of cells: a hypothesis for the evolutionary transitions from abiotic geochemistry to chemoautotrophic prokaryotes, and from prokaryotes to nucleated cells. Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences, 358(1429), 59–85.

Martin, W. F., Sousa, F. L., & Lane, N. (2014). Energy at life's origin. Science, 344(6188), 1092–1093.

Martin, W. F., Weiss, M. C., Neukirchen, S., Nelson-Sathi, S., & Sousa, F. L. (2016). Physiology, phylogeny, and LUCA. Microbial Cell, 3(12), 582.

Mayr, E. (1942). Systematics and the origin of species. Columbia University Press, New York.

McKay et al. (2019). Co-occurring genomic capacity for anaerobic methane and dissimilatory sulfur metabolisms discovered in the Korarchaeota. Nature Microbiology, 4(4), 614–622.

Mereschkowsky, C. (1910). Theorie der zwei Plasmaarten als Grundlage der Symbiogenesis, einer neuen Lehre von der Entstehung der Organismen. Biologisches Centralblatt, 30, 278–288; 289–303; 321–347; 353–367.

Mora, C., Tittensor, D. P., Adl, S., Simpson, A. G., & Worm, B. (2011). How many species are there on Earth and in the ocean? PLoS Biology, 9(8), e1001127.

Moran, N. A. (1996). Accelerated evolution and Muller's rachet in endosymbiotic bacteria. Proceedings of the National Academy of Sciences of the United States of America, 93(7), 2873–2878.

Muller, H. J. (1932). Some genetic aspects of sex. The American Naturalist, 66(703), 118–138.

Munson, M. A., Nedwell, D. B., & Embley, T. M. (1997). Phylogenetic diversity of Archaea in sediment samples from a coastal salt marsh. Applied and Environmental Microbiology, 63(12), 4729–4733.

Muñoz-Gómez, S. A., Kreutz, M., & Hess, S. (2021). A microbial eukaryote with a unique combination of purple bacteria and green algae as endosymbionts. Science Advances, 7(24), eabg4102.

Muñoz-Gómez, S. A. et al. (2022). Site-and-branch-heterogeneous analyses of an expanded dataset favour mitochondria as sister to known Alphaproteobacteria. Nature Ecology & Evolution, 6, 253–262.

N

Nagies, F. S., Brueckner, J., Tria, F. D., & Martin, W. F. (2020). A spectrum of verticality across genes. PLoS Genetics, 16(11), e1009200.

Neidhardt, F. C., Ingraham, J. L., & Schaechter, M. (1990). Physiology of the Bacterial Cell. Sunderland, MA.

Nelson-Sathi, S. et al. (2012). Acquisition of 1,000 eubacterial genes physiologically transformed a methanogen at the origin of Haloarchaea. Proceedings of the National Academy of Sciences of the United States of America, 109(50), 20537–20542.

Naor, A., & Gophna, U. (2013). Cell fusion and hybrids in Archaea: prospects for genome shuffling and accelerated strain development for biotechnology. Bioengineered, 4(3), 126–129.

Newman, D. et al. (2019). Reconstructing and analysing the genome of the Last Eukaryote Common Ancestor to better understand the transition from FECA to LECA. bioRxiv, 538264.

Nielsen, C. (2012). Animal evolution: interrelationships of the living phyla. Oxford University Press on Demand.

Nogales, E. (2001). Structural insights into microtubule function. Annual Review of Biophysics and Biomolecular Structure, 30(1), 397–420.

Nosenko, T. et al. (2013). Deep metazoan phylogeny: when different genes tell different stories. Molecular Phylogenetics and Evolution, 67(1), 223–233.

O

Ochman, H., Lawrence, J. G., & Groisman, E. A. (2000). Lateral gene transfer and the nature of bacterial innovation. Nature, 405(6784), 299–304.

O'Malley, M. A., Leger, M. M., Wideman, J. G., & Ruiz-Trillo, I. (2019). Concepts of the last eukaryotic common ancestor. Nature Ecology & Evolution, 3(3), 338–344.

Опарин А. И. [Oparin A. I.] (1941). Возникновение жизни на Земле [Voznikovyeniye zhizni na Zyemlye // The origin of life on Earth]. 2-е изд., значительно дополненное [2-e izd. znachityelyno dopolnyennoye // second, amended edition]. М.—Л.: Издательство Академии наук СССР [M.-L.: Izdatyelystvo Akademii nauk SSSR // Published by the USSR Academy of Sciences]. 267 pp.

P

Pánek, T. et al. (2015). Combined culture-based and culture-independent approaches provide insights into diversity of jakobids, an extremely plesiomorphic eukaryotic lineage. Frontiers in microbiology, 6, 1288.

Partida-Martinez, L. P. et al. (2007). Endosymbiont-dependent host reproduction maintains bacterial-fungal mutualism. Current Biology, 17(9), 773–777.

Parfrey, L. W. et al. (2010). Broadly sampled multigene analyses yield a well-resolved eukaryotic tree of life. Systematic Biology, 59(5), 518–533.

Park, J. S., & Simpson, A. G. B. (2015). Diversity of heterotrophic protists from extremely hypersaline habitats. Protist, 166, 422–437.

Petit, D., Teppa, E., Cenci, U., Ball, S., & Harduin-Lepers, A. (2018). Reconstruction of the sialylation pathway in the ancestor of eukaryotes. Scientific Reports, 8(1), 1–13.

Pikuta, E. V., Hoover, R. B., & Tang, J. (2007). Microbial extremophiles at the limits of life. Critical Reviews in Microbiology, 33(3), 183–209.

Plutzer, J., Ongerth, J., & Karanis, P. (2010). *Giardia* taxonomy, phylogeny and epidemiology: Facts and open questions. International Journal of Hygiene and Environmental Health, 213(5), 321–333.

Ponce-Toledo, R. I., Deschamps, P., López-García, P., Zivanovic, Y., Benzerara, K., & Moreira, D. (2017). An early-branching freshwater cyanobacterium at the origin of plastids. Current Biology, 27(3), 386–391.

Porter, S. M. (2020). Insights into eukaryogenesis from the fossil record. Interface Focus, 10(4), 20190105.

Preiner, M. et al. (2018). Serpentinization: connecting geochemistry, ancient metabolism and industrial hydrogenation. Life, 8(4), 41.

R

Radzvilavicius, A. L. (2016). Mitochondrial genome erosion and the evolution of sex. Bioessays, 38(10).

Radzvilavicius, A. L., & Blackstone, N. W. (2015). Conflict and cooperation in eukaryogenesis: implications for the timing of endosymbiosis and the evolution of sex. Journal of the Royal Society Interface, 12(111), 20150584.

Radzvilavicius, A. L., Hadjivasiliou, Z., Pomiankowski, A., & Lane, N. (2016). Selection for mitochondrial quality drives evolution of the germline. PLoS Biology, 14(12), e2000410.

Raff, R. A., & Mahler, H. R. (1972). The non symbiotic origin of mitochondria. Science, 177(4049), 575–582.

Riisberg et al. (2009). Seven gene phylogeny of heterokonts. Protist, 160, 191–204.

Robinson, K. M., Sieber, K. B., & Dunning Hotopp, J. C. (2013). A review of bacteria-animal lateral gene transfer may inform our understanding of diseases like cancer. PLoS Genetics, 9(10), e1003877.

Roger, A. J., Muñoz-Gómez, S. A., & Kamikawa, R. (2017). The origin and diversification of mitochondria. Current Biology, 27(21), R1177–R1192.

Roger, A. J., & Simpson, A. G. (2009). Evolution: revisiting the root of the eukaryote tree. Current Biology, 19(4), R165–R167.

Roger, A. J., Susko, E., & Leger, M. M. (2021). Evolution: Reconstructing the Timeline of Eukaryogenesis. Current Biology, 31(4), R193–R196.

Russell, M. J., & Hall, A. J. (1997). The emergence of life from iron monosulphide bubbles at a submarine hydrothermal redox and pH front. Journal of the Geological Society, 154(3), 377–402.

S

Say, R. F., & Fuchs, G. (2010). Fructose 1, 6-bisphosphate aldolase/phosphatase may be an ancestral gluconeogenic enzyme. Nature, 464(7291), 1077–1081.

Schleper, C., & Sousa, F. L. (2020). Meet the relatives of our cellular ancestor. Nature 577, 478–479.

Searcy D, G. (1992). Origins of mitochondria and chloroplasts from sulphur-based symbioses. In: Hartman, H., & Matsuno, K. (Eds.) The origin and evolution of the cell, pp. 47 – 78. Singapore: World Scientific.

Shen, H. et al. (2018). Large-scale phylogenomic analysis resolves a backbone phylogeny in ferns. GigaScience, *7*(2), gix116.

Shiratori, T., Nakayama, T., & Ishida, K. (2015). A new deep-branching stramenopile, *Platysulcus tardus* gen. nov., sp. nov. Protist, 166, 337–348.

Sierra, R. et al. (2013). Deep relationships of Rhizaria revealed by phylogenomics: A farewell to Haeckel's Radiolaria. Molecular Phylogenetics and Evolution, 67, 53–59.

Sierra, R. et al. (2016). Evolutionary origins of rhizarian parasites. Molecular Biology and Evolution, 33, 980–983.

Simion, P. et al. (2017). A large and consistent phylogenomic dataset supports sponges as the sister group to all other animals. Current Biology, 27(7), 958–967.

Simpson, A. G. B., Slamovits, C. H., & Archibald, J. M. (2017). Protist diversity and eukaryote phylogeny. Handbook of the Protists, 1–21.

Skejo, Josip et al. (2021). **Evidence for a syncytial origin of eukaryotes from ancestral state reconstruction**. Genome Biology and Evolution, 13(7), evab096, 1–14. (PAPER 3 of the thesis)

Skejo, Josip, & Franjević, D. (2020). **Eukaryotes are a holophyletic group of polyphyletic origin**. Frontiers in Microbiology, 11(1380), 1–6. (PAPER 1 of the thesis)

Smith, J. E. (1987). Erythrocyte membrane: structure, function, and pathophysiology. Veterinary Pathology, 24(6), 471–476.

Sokol, N. W. et al. (2022). Life and death in the soil microbiome: how ecological processes influence biogeochemistry. Nature Reviews Microbiology, 1-16 (online ahead of print), doi: 10.1038/s41579-022-00695-z.

Söllinger, A., & Urich, T. (2019). Methylotrophic methanogens everywhere—physiology and ecology of novel players in global methane cycling. Biochemical Society Transactions, 47(6), 1895–1907.

Spain, A. M., Krumholz, L. R., & Elshahed, M. S. (2009). Abundance, composition, diversity and novelty of soil Proteobacteria. The ISME Journal, 3(8), 992–1000.

Spatafora, J. W. et al. (2016). A phylum-level phylogenetic classification of zygomycete fungi based on genome-scale data. Mycologia, 108(5), 1028–1046.

Speijer, D., Lukeš, J., & Eliáš, M. (2015). Sex is a ubiquitous, ancient, and inherent attribute of eukaryotic life. Proceedings of the National Academy of Sciences of the United States of America, 112(29), 8827–8834.

Sprang, A. et al. (2015). Complex archaea that bridge the gap between prpkaryocytes and eukaryocytes. Nature, 521, 173–179.

Stechmann, A., & Cavalier-Smith, T. (2002). Rooting the eukaryote tree by using a derived gene fusion. Science, 297(5578), 89–91.

Steiner, J. M. (2010) Technical notes: Growth of *Cyanophora paradoxa*. Journal of Endocytobiosis & Cell Research, 20, 62–67.

Stolz, J. F. (2000). Structure of microbial mats and biofilms. In Microbial sediments (pp. 1–8). Springer, Berlin, Heidelberg.

Strassert, J. F. et al. (2016). Moramonas marocensis gen. nov., sp. nov.: a jakobid flagellate isolated from desert soil with a bacteria-like, but bloated mitochondrial genome. Open Biology, 6(2), 150239.

Strassert, J. F. et al. (2019). New phylogenomic analysis of the enigmatic phylum Telonemia further resolves the eukaryote tree of life. Molecular Biology and Evolution, 36(4), 757–765.

Subedi, B. P. et al. (2021). Archaeal pseudomurein and bacterial murein cell wall biosynthesis share a common evolutionary ancestry. FEMS Microbes, 2.

Swart, E. C. et al. (2013). The Oxytricha trifallax macronuclear genome: a complex eukaryotic genome with 16,000 tiny chromosomes. PLoS Biology, 11(1), e1001473.

T

Tekle, Y. I. et al. (2017). Amoebozoans are secretly but ancestrally sexual: evidence for sex genes and potential novel crossover pathways in diverse groups of amoebae. Genome Biology and Evolution, 9(2), 375–387.

Thiergart, T., Landan, G., Schenk, M., Dagan, T., & Martin, W. F. (2012). An evolutionary network of genes present in the eukaryote common ancestor polls genomes on eukaryotic and mitochondrial origin. Genome Biology and Evolution, 4(4), 466–485.

Tikhonenkov, D. V. (2020). Predatory flagellates-the new recently discovered deep branches of the eukaryotic tree and their evolutionary and ecological significance. Protistology, 14(1), 15–22.

Tikhonenkov, D. V. et al. (2014). Description of *Colponema vietnamica* sp.n. and *Acavomonas peruviana* n. gen. n. sp., two new alveolate phyla (Colponemidia nom. nov. and Acavomonidia nom. nov.) and their contributions to reconstructing the ancestral state of alveolates and eukaryotes. PloS One, 16, e95467.

Timmis, J. N., Ayliffe, M. A., Huang, C. Y., & Martin, W. (2004). Endosymbiotic gene transfer: organelle genomes forge eukaryotic chromosomes. Nature Reviews Genetics, 5(2), 123–135.

Tria, F. D., Brückner, J., **Skejo, Josip**, Xavier, J. C., Kapust, N., Knopp, M., Wimmer, J. L. E., Nagies, F. S. P., Zimorski, V., Gould, S. B., Garg, S., & Martin, W. F. et al. (2021). **Gene duplications trace mitochondria to the onset of eukaryote complexity**. Genome Biology and Evolution, 13(5), evab055, 1–17. (PAPER 2 of the thesis)

Tromer, E. C., van Hooff, J. J., Kops, G. J., & Snel, B. (2019). Mosaic origin of the eukaryotic kinetochore. Proceedings of the National Academy of Sciences of the United States of America, 116(26), 12873–12882.

Turland, N. J. et al. (2018). International Code of Nomenclature for algae, fungi, and plants (Shenzhen Code) adopted by the Nineteenth International Botanical Congress Shenzhen, China, July 2017. Koeltz Botanical Books.

U

Ultsch, G. R. (1996). Gas exchange, hypercarbia and acid-base balance, paleoecology, and the evolutionary transition from water-breathing to air-breathing among vertebrates. Palaeogeography, Palaeoclimatology, Palaeoecology, 123(1–4), 1–27.

V

Valero, M., Richerd, S., Perrot, V., & Destombe, C. (1992). Evolution of alternation of haploid and diploid phases in life cycles. Trends in Ecology & Evolution, 7(1), 25–29.

Vellai, T., & Vida, G. (1999). The origin of eukaryotes: the difference between prokaryotic and eukaryotic cells. Proceedings of the Royal Society of London. Series B: Biological Sciences, 266(1428), 1571–1577.

Vincent, L., Colón-Santos, S., Cleaves, H. J., Baum, D. A., & Maurer, S. E. (2021). The prebiotic kitchen: A guide to composing prebiotic soup recipes to test origins of life hypotheses. Life, 11(11), 1221.

Volkova, E., & Kudryavtsev, A. (2017). Description of Neoparamoeba longipodia n. sp. and a new strain of Neoparamoeba aestuarina (Page, 1970)(Amoebozoa, Dactylopodida) from deep-sea habitats. European Journal of Protistology, 61, 107–121.

W

Weiss, M. C., Preiner, M., Xavier, J. C., Zimorski, V., & Martin, W. F. (2018). The last universal common ancestor between ancient Earth chemistry and the onset of genetics. PLoS Genetics, 14(8), e1007518.

Weiss, M. C., Sousa, F. L., Mrnjavac, N., Neukirchen, S., Roettger, M., Nelson-Sathi, S., & Martin, W. F. (2016). The physiology and habitat of the last universal common ancestor. Nature Microbiology, 1(9), 1–8.

Whelan, N. V., Kocot, K. M., Moroz, L. L., & Halanych, K. M. (2015). Error, signal, and the placement of Ctenophora sister to all other animals. Proceedings of the National Academy of Sciences of the United States of America, 112(18), 5773–5778.

Whittaker, R. H. (1969). New concepts of kingdoms of organism. Science, 163, 150–159.

Wickett et al. (2014). Phylotranscriptomic analysis of the origin and early diversification of land plants. Proceedings of the National Academy of Sciences of the United States of America, 111, 4859–4868.

Wilkins, J. S. (2018). Species: the evolution of the idea. CRC Press.

Williams, T. A. et al. (2017). Integrative modeling of gene and genome evolution roots the archaeal tree of life. Proceedings of the National Academy of Sciences of the United States of America, 114(23), E4602–E4611.

Williams, T. A. et al. (2020). Phylogenomics provides robust support for a two-domains tree of life. Nature Ecology & Evolution, 4(1), 138–147.

Willis, K., & McElwain, J. (2014). The evolution of plants. Oxford University Press.

Woese, C. R., Kandler, O., & Wheelis, M. L. (1990). Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya. Proceedings of the National Academy of Sciences of the United States of America, 87(12), 4576–4579.

Wu, B. et al. (2019). Current insights into fungal species diversity and perspective on naming the environmental DNA sequences of fungi. Mycology, 10(3), 127–140.

X

Xavier, J. C., Gerhards, R. E., Wimmer, J. L., Brueckner, J., Tria, F. D., & Martin, W. F. (2021). The metabolic network of the last bacterial common ancestor. Communications Biology, 4(1), 1–10.

Y

Yabuki, A. et al. (2013). Fine structure of *Telonema subtilis* Griessmann, 1913: a flagellate with a unique cytoskeletal structure among eukaryotes. Protist, 164(4), 556–569.

Yabuki, A. et al. (2014). *Palpitomonas bilix* represents a basal cryptist lineage: Insight into the character evolution in Cryptista. Scientific Reports, 4, 4641.

Yubuki, N. et al. (2015). Morphological identities of two different marine stramenopile environmental sequence clades: *Bicosoeca kenaiensis* (Hilliard, 1971) and *Cantina marsupialis* (Larsen and Patterson, 1990) gen. nov., comb. nov. Journal of Eukaryotic Microbiology, 62, 532–542.

Z

Zardoya, R., Cao, Y., Hasegawa, M., & Meyer, A. (1998). Searching for the closest living relative (s) of tetrapods through evolutionary analyses of mitochondrial and nuclear data. Molecular Biology and Evolution, 15(5), 506–517.

Zaremba-Niedzwiedzka, K. et al. (2017). Asgard archaea illuminate the origin of eukaryotic cellular complexity. Nature, 541(7637), 353–358.

Zhang, N., Luo, J., & Bhattacharya, D. (2017). Advances in fungal phylogenomics and their impact on fungal systematics. Advances in genetics 100, 309–328.

Zhang, Z., Wu, Y., & Zhang, X. H. (2018). Cultivation of microbes from the deep-sea environments. Deep Sea Research Part II: Topical Studies in Oceanography, 155, 34–43.

Zhao, S. et al. (2012). Collodictyon—an ancient lineage in the tree of eukaryotes. Molecular Biology and Evolution, 29(6), 1557–1568.

Zhu, Q. et al. (2019). Phylogenomics of 10,575 genomes reveals evolutionary proximity between domains Bacteria and Archaea. Nature Communications, 10(1), 1–14.

Zimorski, V., Ku, C., Martin, W. F., & Gould, S. B. (2014). Endosymbiotic theory for organelle origins. Current Opinion in Microbiology, 22, 38–48.

Zimorski, V., Mentel, M., Tielens, A. G., & Martin, W. F. (2019). Energy metabolism in anaerobic eukaryotes and Earth's late oxygenation. Free Radical Biology and Medicine, 140, 279–294.

Zmasek, C. M., & Godzik, A. (2011). Strong functional patterns in the evolution of eukaryotic genomes revealed by the reconstruction of ancestral protein domain repertoires. Genome Biology, 12(1), 1–13.

AUTHOR'S CURRICULUM VITAE

Josip Skejo
Doctoral candidate in Biology,
Master of Experimental Biology (Zoology),
Bachelor of Biology

Birth	29th November 1992 in Šibenik, Croatia
Parents	Dubravka Skejo r. Bilić, Knin & Marko Skejo, Sedramić
Evolutionary origin	mtDNA hg W1e, yDNA hg I-S17250 (within I2a-L621)
Blood type	A Rh^+
Private address	Alagovićeva 21, Zagreb, Croatia
Private website	https://linktr.ee/skejo
Work	University of Zagreb, Faculty of Science, Department of Biology, Division of Zoology, Evolution Lab
Work address	Rooseveltov trg 6, Zagreb, Croatia
Work website	https://www.pmf.unizg.hr/biol/en/josip.skejo
Email	skejo.josip@gmail.com, jskejo@biol.pmf.hr
About me	24/7 naturalist, zoologist, field biologist, and wet-lab biologist, with more than 50 publications, including two monographs. Primarily working on taxonomy, biogeography and ecology of grasshoppers and bush-crickets (Orthoptera) of Europe, with special emphasis on the Dinaric Alps; as well as on pygmy grasshoppers (Tetrigidae) of the whole world with special emphasis on taxa from South-eastern Asia. I have hitherto described more than 40 species new to science, and also some genera and higher taxa. Interested in conservation, phylogeny, cladistics, evolution, botany and various other fields within biology, curiosity drives me. For me, this five-years of the PhD represented the first major expedition to the world of prokaryotes and protists.

Languages Croatian is my mother tongue with native dialect Dinaric Stokavian Ikavian (or younger Ikavian, close to Eastern Herzegovinian, from which official norms of Croatian, Serbian, Bosnian, and Montenegrin originated); I am proficient in English; good in many Slavic languages (e.g., Russian, Ukrainian, Serbian, Old Church Slavonic); proficient in classical tongues (Latin, Ancient Greek); and have basic knowledge of many other languages (Romance, such as French and Spanish; Germanic, such as German; but also Mandarin; and Korean etc.).

Education

2017–2022 PhD in biology, Department of Biology, Faculty of Science, University of Zagreb

Doctoral thesis (2022). Reconstruction of the Last Eukaryotic Common Ancestor by cladistic and phylogenetic approach.

2013–2017 Master's course of experimental biology, module Zoology, Department of Biology, Faculty of Science, University of Zagreb

Master thesis (2017). Taxonomic revision of the pygmy devils (Tetrigidae: Discotettiginae) with online social media as a new tool for discovering hidden diversity. 235 pp.

2011–2013 Undergraduate biology course, Department of Biology, Faculty of Science, University of Zagreb

Bachelor thesis (2013). Taxonomy and distribution of the Croatian groundhoppers (Orthoptera: Tetrigidae). 45 pp. [in Croatian]

2007–2011 Franciscan Grammar School, Sinj, Croatia

1999–2007 Primary School, Šibenik, Croatia

Work experience

2017–today Assistant, Department of Biology, Faculty of Science, University of Zagreb

Publications

Books

Hochkirch, A., Nieto, A., García Criado, M., Cálix, M., Braud, Y., Buzzetti, F., Chobanov, D., Odé, B., Presa Asensio, J., Willemse, L., Zuna-Kratky, T., Barranco Vega, P., Bushell, M., Clemente, M., Correas, J., Dusoulier, F., Ferreira, S., Fontana, P., García, M., Heller, K., Iorgu, I., Ivković, S., Kati, V., Kleukers, R., Krištín, A., Lemonnier-Darcemont, M., Lemos, P., Massa, B., Monnerat, C., Papapavlou, K., Prunier, F., Pushkar, T., Roesti, C., Rutschmann, F., Şirin, D., Skejo, J., Szövényi, G., Tzirkalli, E., Vedenina, V., Domenech, J., Barros, F., Cordero Tapia, P., Defaut, B., Fartmann, T., Gomboc, S., Gutiérrez-Rodríguez, J., Holuša, J., Illich, I., Karjalainen, S., Kočárek, P., Korsunovskaya, O., Liana, A., López, H., Morin, D., Olmo-Vidal, J., Puskás, G., Savitsky, V., Thomas Stalling & Tumbrinck, J. (2016) *European Red List of Grasshoppers, Crickets and Bush-crickets*. Luxembourg, European Commission, 87 pp.

Muhammad, A., Tan, M., Abdullah Nurul Ashikin, Azirun, M., Bhaskar, D. & Skejo, J. (2018) *An annotated catalogue of the pygmy grasshoppers of the tribe Scelimenini Bolívar, 1887 (Orthoptera: Tetrigidae) with two new* Scelimena *species from the Malay Peninsula and Sumatra*. Zootaxa, 4485(1), 70 pp.

Ozimec, R., Baković, N., Baričević, L., Božić, B., Drakšić, M., Ernoić, M., Fressel, N., Kučinić, M., Kušan, I., Lacković, D., Martinko, M., Matočec, N., Samardžić, M., Skejo, J. & Šincek, D. (2016) *Đurđevec Sands.* ADIPA - Društvo za istraživanje i očuvanje prirodoslovne raznolikosti Hrvatske, Zagreb, 96 pp. [in Croatian]

Skejo, J., Rebrina, F., Szövényi, G., Puskás, G. & Tvrtković, N. (2018) *The first annotated checklist of Croatian crickets and grasshoppers (Orthoptera: Ensifera, Caelifera).* Zootaxa, 4533(1), 95 pp.

Book chapters

Adžić, K., Deranja, M., Pavlović, M., Tumbrinck, J. & Skejo, J. (2021) Endangered Pygmy Grasshoppers (Tetrigidae). In: DellaSala, D. & Goldstein, M. (ed.) *Imperiled: The Encyclopaedia of Conservation*. Oxford, Elsevier, pp. 1–11.

Skejo, J. & Rebrina, F. (2020) Croatia. In: Kleukers, R. & Felix, R. (ed.) *Grasshopper conservation in Europe.* Leiden, EIS Kenniscentrum Insecten & Naturalis Biodiversity Center, pp. 19–20.

Tumbrinck, J., Skejo, J. (2017) Taxonomic and biogeographic revision of the New Guinean genus *Ophiotettix* Walker, 1871 (Tetrigidae: Metrodorinae: Ophiotettigini trib. nov.), with the descriptions of 33 new species. In: Telnov, D., Barclay, M. & Pauwels, O. (ed.) *Biodiversity, biogeography and nature conservation in Wallacea and New Guinea* (Volume III). Riga, Latvia, The Entomological Society of Latvia, pp. 525–580.

Educational material

Blažetić, S., Heffer, M., Lucić, A., Sedlar, Z., Skejo, J., Bendelja, D., Lukša, Ž. (2019) *Biologija 2: udžbenik za eksperimentalni program biologije u drugom razredu gimnazije// Biology 2 - A textbook for an experimental biology program, 2nd grade Gymnasium*. Zagreb. Školska knjiga, 376 pp. [in Croatian]

Lucić, A., Skejo, J., Heffer, M., Sedlar, Z., Blažetić, S., Bendelja, D., Lukša, Ž. (2020) Biologija 2 – udžbenik biologije s dodatnim digitalnim sadržajima u drugom razredu gimnazije. // *Biology 2 – A textbook for 2nd grade Gymnasium biology program, with additional digital materials.* Zagreb. Školska knjiga, 192 pp. [in Croatian]

Journal articles

2022

Martinović, M., Ćato, S., Lengar, M., Skejo, J. (2022) First records of three exotic giant mantid species on the Croatian coast. *Journal of Orthoptera Research*. Accepted (https://jor.pensoft.net/article/76075/)

2021

Hlebec, D., Sivec, I., Podnar, M., Skejo, J. & Kučinić, M. (2021) Morphological and molecular characterisation of the Popijač's Yellow Sally, *Isoperla popijaci* sp. nov., a new stenoendemic stonefly species from Croatia (Plecoptera, Perlodidae). *ZooKeys*, 1078, 85–106.

Kasalo, N., Deranja, M., Adžić, K., Sindaco, R. & Skejo, J. (2021) Discovering insect species based on photographs only: The case of a nameless species of the genus *Scaria* (Orthoptera: Tetrigidae). *Journal of Orthoptera Research*, 30 (2), 173–184.

Kranjec Orlović, J., Bulovec, I., Franjević, M., Franjević, D., Skejo, J., Biliškov, M., Diminić, D. & Hrašovec, B. (2021) Preliminary results on narrow-leaved ash (*Fraxinus angustifolia* Vahl) and green ash (*Fraxinus pennsylvanica* Marshall) seed entomofauna in Croatia. *Šumarski list,* 145 (3-4), 147–154.

Mathieu, É., Pavlović, M. & Skejo, J. (2021) The true colours of the Formidable Pygmy Grasshopper (*Notocerus formidabilis* Günther, 1974) from the Sava region (Madagascar). *ZooKeys*, 1042, 41–50.

Patano, R., Mohagan, A., Tumbrinck, J., Amoroso, V. & Skejo, J. (2021) Horned and spiky: *Tegotettix derijei* sp. n. (Orthoptera: Tetrigidae) is a peculiar new pygmy grasshopper species from Mindanao. *Zootaxa*, 4933 (2), 198–210.

Skejo, J., Garg, S., Gould, S., Hendriksen, M., Tria, F., Bremer, N., Franjević, D., Blackstone, N. & Martin, W. (2021) Evidence for a syncytial origin of eukaryotes from ancestral state reconstruction. Genome *Biology and Evolution*, 13 (7), 1–14.

Tria, F., Brueckner, J., Skejo, J., Xavier, J., Kapust, N., Knopp, M., Wimmer, J., Nagies, F., Zimorski, V., Gould, S., Garg, S. & Martin, W. (2021) Gene duplications trace mitochondria to the onset of eukaryote complexity. *Genome Biology and Evolution*, 13 (5), 1–17.

Tumbrinck, J., Deranja, M., Adžić, K., Pavlović, M. & Skejo, J. (2021) Erratum: Josef Tumbrinck, Maks Deranja, Karmela Adžić, Marko Pavlović, Josip Skejo (2020) Cockscomb-shaped twighopper, *Cladonotus bhaskari* sp. n., a new and rare pygmy grasshopper species from Sri Lanka (Orthoptera: Tetrigidae: Cladonotinae). Zootaxa, 4821: 333–342. *Zootaxa*, 4951 (3), 598-598.

Zha, L., Skejo, J., Mao, B. & Ding, J. (2021) Taxonomic revision of *Phaesticus* Uvarov and synonymy with *Flatocerus* Liang & Zheng syn. nov. (Orthoptera: Tetrigidae). *Zootaxa*, 4965 (3), 501–514.

2020

Adžić, K., Deranja, M., Franjević, D. & Skejo, J. (2020) Are Scelimeninae (Orthoptera: Tetrigidae) monophyletic and why it remains a question? *Entomological news*, 129 (2), 128–146.

Bhaskar, D., Stermšek, S., Easa, P., Franjević, D. & Skejo, J. (2020) Wide-nosed pygmy grasshoppers (Cladonotinae: Cladonotini, Xerophyllini) of India and Sri Lanka: catalogue with an identification key and description of a new species of the genus *Tettilobus*. *Zootaxa*, 4894 (3), 474–500.

Felix, R., Heller, K., Odé, B., Rebrina, F. & Skejo, J. (2020) Island mysteries in the spotlight: *Barbitistes kaltenbachi* and *Rhacocleis buchichii*, the only bush-cricket species endemic to Croatia (Orthoptera, Tettigoniidae). *ZooKeys*, 936, 25–60.

Ivković, S. & Skejo, J. (2020) Who is jumping in a Serbian bog? – Orthopteran fauna of the Vlasina region. *Travaux du Muséum National d'Histoire Naturelle "Grigore Antipa",* 63 (2), 141–160.

Pavlović, M., Sule, D. & Skejo, J. (2020) Pregled skakavaca i zrikavaca (Insecta: Orthoptera) Šolte i usporedba s faunom Brača. // Overview of the grasshoppers and crickets (Insecta: Orthoptera) of Šolta and comparison with Brač. *Bašćina (Grohote)*, 29, 120–134. [in Croatian]

Skejo, J., Connors, M., Hendriksen, M., Lambert, N., Chong, G., McMaster, I., Monaghan, N., Rentz, D., Richter, R., Rose, K. & Franjević, D. (2020) Online social media tells a story of *Anaselina*, *Paraselina*, and *Selivinga* (Orthoptera, Tetrigidae), rare Australian pygmy grasshoppers. *ZooKeys*, 948, 107–119.

Skejo, J., Deranja, M. & Adžić, K. (2020) Pygmy hunchback of New Caledonia: *Notredamia dora* gen. n. et sp. n. - a new cladonotin (Caelifera: Tetrigidae) genus and species from Oceania. *Entomological news*, 129 (2), 170–185.

Skejo, J. & Franjević, D. (2020) Eukaryotes are a holophyletic group of polyphyletic origin. *Frontiers in microbiology*, 11, 1380, 6.

Skejo, J., Medak, K., Pavlović, M., Kitonić, D., Miko, J. & Franjević, D. (2020) The story of the Malagasy devils (Orthoptera, Tetrigidae): *Holocerus lucifer* in the north and *H. devriesei* sp. nov. in the south? *ZooKeys*, 957, 1–15.

Tumbrinck, J., Deranja, M., Adžić, K., Pavlović, M. & Skejo, J. (2020) Cockscomb-shaped twighopper, *Cladonotus bhaskari* sp. n., a new and rare pygmy grasshopper species from Sri Lanka (Orthoptera: Tetrigidae: Cladonotinae). *Zootaxa*, 4821 (2), 333–342.

2019

Bhaskar, D., Easa, P. S., Sreejith, K. A., Skejo, J., & Hochkirch, A. (2019). Large scale burning for a threatened ungulate in a biodiversity hotspot is detrimental for grasshoppers (Orthoptera: Caelifera). *Biodiversity and Conservation*, 28 (12), 3221–3237.

Đukić, A., Mirić, R., Skejo, J., Rajkov, S. & Tot, I. (2019) Survey on the damselfly and dragonfly fauna (Insecta: Odonata) of the Landscape of outstanding features "Vlasina". *Kragujevac Journal of Science*, 41, 133–146.

Kehoe, L., Reis, T., Virah-Sawmy, M., Balmford, A., Kuemmerle, T. & 604 signatories (2019) Make EU trade with Brazil sustainable. *Science*, 364 (6438), 341–341.

Škorput, J., Novak Morić, A., Martinović, M., van der Heyden, T. & Skejo, J. (2019) *Solenosthedium bilunatum* (Heteroptera: Scutelleridae) at the Adriatic Coast of Croatia. *Entomologie heute*, 31, 25–29.

Skejo, J., Gupta, S., Chandra, K., Panhwar, W. & Franjević, D. (2019) Oriental macropterous leaf-mimic pygmy grasshoppers—genera *Oxyphyllum* and *Paraphyllum* (Orthoptera: Tetrigidae) and their taxonomic assignment. *Zootaxa*, 4590 (5), 546–560.

Thomas, M. J., Skejo, J., & Heads, S. W. (2019). The last batrachideine of Europe: A new genus and species of pygmy grasshopper (Orthoptera: Tetrigidae) from Eocene Baltic amber. *Zootaxa*, 4686(3), 435–445.

2018

Szövényi, G., Skejo, J., Rebrina, F., Tvrtković, N. & Puskás, G. (2018) First data on the Orthoptera diversity of Poštak Mountain and its surroundings (Croatia). *Annales de la Société entomologique de France*, 54 (6), 1–11.

Skejo, J., Gupta, S. & Tumbrinck, J. (2018) Nymph inadvertently described as new species for a fourth time? On the identity of *Euscelimena hardi* (Tetrigidae: Scelimeninae) with general remarks on the identification of pygmy grasshopper nymphs. *Zootaxa*, 4418 (1), 93–97.

Ivković, S., Pantović, U., & Skejo, J. (2018). Ovčar–Kablar Gorge (SW Serbia)–a new hotspot of Orthoptera diversity. *Annales de la Société entomologique de France*, 54 (3), 257–272.

2017

Skejo, J. & Bertner, P. (2017) No more dust and exoskeletons – *in vivo* photographic records provide new data on *Eufalconius pendleburyi* Günther, 1938 (Orthoptera: Tetrigidae) from the Titiwangsa Mts. *Annales zoologici*, 67 (4), 665–672.

Silva, D., Skejo, J., Pereira, M., De Domenico, F. & Sperber, C. (2017) Comments on the recent changes in taxonomy of pygmy unicorns, with description of a new species of *Metopomystrum* from Brazil (Insecta, Tetrigidae, Cleostratini, Miriatrini). *ZooKeys*, 702 (9), 1–18.

Lehmann, A., Devriese, H., Tumbrinck, J., Skejo, J., Lehmann, G. & Hochkirch, A. (2017) The importance of validated alpha taxonomy for phylogenetic and DNA barcoding studies: a comment on species identification of pygmy grasshoppers (Orthoptera, Tetrigidae). *ZooKeys*, 679, 139–144.

2016

Skejo, J. (2016) On the taxonomy of the genus *Rosacris* Bolívar, 1931 (Orthoptera: Tetrigidae). *Entomologie heute*, 28, 43–52.

Skejo, J. & Caballero, J. (2016) A hidden pygmy devil from the Philippines: *Arulenus miae* sp. nov. – a new species serendipitously discovered in an amateur Facebook post (Tetrigidae: Discotettiginae). *Zootaxa*, 4067 (3), 383–393.

Skejo, J. & Pantović, U. [Скејо, Ј. & Пантовић, У.] (2016) Први преглед фауне правокрилаца (Insecta: Polyneoptera: Orthoptera) Овчарско–кабларске клисуре (Западна Србија). // The first overview of the grasshoppers' fauna (Insecta: Polyneoptera: Orthoptera) of the Ovčar–Kablar Gorge (West Serbia). *Бележник Овчарско–кабларске клисуре*, 6 (1), 30–37. [in Serbian]

2015

Kaya, S., Chobanov, D., Skejo, J., Heller, K. & Çiplak, B. (2015) The Balkan *Psorodonotus* (Orthoptera: Tettigoniidae): Testing the existing taxa confirmed presence of three distinct species. *European journal of entomology*, 112 (3), 525–541.

Rebrina, F., Skejo, J., Lucić, A. & Hudina, S. (2015) Trait variability of the signal crayfish (*Pacifastacus leniusculus*) in a recently invaded region reflects potential benefits and trade-offs during dispersal. *Aquatic invasions*, 10 (1), 41–50.

Rebrina, F., Skejo, J. & Tvrtković, N. (2015) First results of inventarisation of Blattodea, Mantodea and Orthoptera (Insecta: Polyneoptera) of the Dinara Mountain area. *Annales de la Societe entomologique de france*, 51 (1), 60–69.

Skejo, J. & Gupta, S. (2015) On the specific status of *Hedotettix cristatus* Karny, 1915 (Tetrigidae: Tetriginae). *Zootaxa*, 4018 (4), 584–592.

Skejo, J. & Ivković, S. (2015) *Chorthippus bornhalmi* in the heart of the Balkans (Acrididae: Gomphocerinae). *Articulata - Zeitschrift der Deutschen Gesellschaft für Orthopterologie*, 30, 81–90.

Skejo, J., Rebrina, F., Tvrtković, N., Gomboc, S., Heller, K.-G. (2015) More than a century old *'Platycleis Kraussi* case' finally resolved (Tettigoniidae: Platycleidini). *Zootaxa*, 3990 (4), 497-524.

Skejo, J. & Sule, D. (2015) Prvi doprinos poznavanju raznolikosti zrikavaca i skakavaca (Insecta: Orthoptera) Šolte. // First contribution to the knowledge of the diversity of grasshoppers and crickets (Insecta: Orthoptera) of the Island of Šolta. *Bašćina (Grohote),* 24, 19–24. [in Croatian]

2014

Rebrina, F., Battiston, R. & Skejo, J. (2014) Are *Empusa pennata* and *Bolivaria brachyptera* really present in Croatia? A reply to Kranjčev (2013) with a critical review of the Mantid taxa found in Croatia. *Entomologia Croatica*, 18 (1–2), 17–25.

Skejo, J., Rebrina, F., Buzzetti, F., Ivković, S., Rašić, A. & Tvrtković, N. (2014) First records of Croatian and Serbian Tetrigidae (Orthoptera: Caelifera) with description of a new subspecies of *Tetrix transsylvanica* (Bazyluk & Kis, 1960). *Zootaxa*, 3856 (3), 419–, 14.

Skejo, J. & Stanković, M. (2014) of the Special Nature Reserve Zasavica (S Vojvodina, Serbia) with special emphasis on *Zeuneriana amplipennis*. *Articulata – Zeitschrift der Deutschen Gesellschaft für Orthopterologie*, 29 (1), 9–20.

Šerić Jelaska, L. & Skejo, J. (2014) Catalogue of the Entomological collections of The Division of Zoology of The Faculty of Science in Zagreb - Collection of Orthoptera (Polyneoptera, Orthopterida) of Boža Pokopac. *Entomologia Croatica*, 18 (1–2), 59–71.

2013

Skejo, J. & Rebrina, F. (2013) *Rammeihippus dinaricus* (Götz, 1970) (Orthoptera: Acrididae) – a new genus and species for the orthopteran fauna of Croatia and the first record of the species since description. *Natura Croatica*, 22 (1), 37–41.

Skejo, J. & Stanković, M. (2013) The westernmost localities for the bush-cricket *Leptophyes discoidalis* (Tettigoniidae: Phaneropterinae). *Natura Croatica*, 22 (2), 339–341.

Conference proceedings papers

Rebrina, F., Skejo, J. & Šerić Jelaska, L. (2017) Zajednice ravnokrilaca (Insecta: Orthoptera) suhih i vlažnih travnjaka na području NP „Krka“. // Orthoptera communities (Insecta: Orthoptera) of the dry and wet grasslands in Krka National Park. In: Marguš, D. (ed.) *Zbornik radova sa znanstveno - stručnog skupa Vizija i izazovi upravljanja zaštićenim područjima prirode u Republici Hrvatskoj: Aktivna zaštita i održivo upravljanje u Nacionalnom parku 'Krka'*. Šibenik, Javna ustanova „Nacionalni park Krka”, pp. 259–268. [in Croatian]

Theses I have mentored

Regul, J. (2022) Sistematika skakavaca veslača (Insecta: Orthoptera: Scelimeninae) temeljena na morfološkim svojstvima. // Systematics of pygmy rowerhoppers (Insecta: Orthoptera: Scelimeninae) based on morphological characters. *Master thesis, Faculty of Science, Zagreb*. [in Croatian]

Adžić, K. (2021) Pygmy Grasshoppers (Orthoptera: Tetrigidae) of Peninsular Malaysia. *Master thesis, Faculty of Science, Zagreb.*

Deranja, M. (2021) Cladistic analysis of diagnostic characters of wide-nosed pygmy grasshoppers (Orthoptera: Tetrigidae: Cladonotinae). *Master thesis, Faculty of Science, Zagreb.*

Kulessa, A. (2021) Evolution of Malagasy pygmy grasshoppers (Orthoptera: Tetrigidae). // Die Evolution der madagassischen Dornschrecken (Orthoptera: Tetrigidae). *Bachelor thesis, Heinrich-Heine-Universität Düsseldorf, Düsseldorf*. [in German]

Pavlović, M. (2020) Biogeografija skakavaca i zrikavaca (Orthoptera: Caelifera, Ensifera) jadranskog područja. // Biogeography of grasshoppers and crickets (Orthoptera: Caelifera, Ensifera) of the Adriatic area. *Master thesis, Faculty of Science, Zagreb*. [in Croatian]

Bremer, N. (2020) Phylogenetic analysis and ancestral state reconstruction of the last eukaryotic common ancestor. *Master thesis, Heinrich-Heine-Universität Düsseldorf, Düsseldorf*.

Raos, D. (2018) Povezanost ovulacijskog ciklusa i razine luteinizirajućeg hormona s preferencijama žena prema muškim partnerima. // The relationship between the ovulation cycle and 17β-estradiol level with the preferences of women to male partners. *Master thesis, Faculty of Science, Zagreb*. [in Croatian]

Šapina, I. (2018) Kladistički pristup klasifikaciji roda *Hirrius* Bolívar, 1887 (Orthoptera, Tetrigidae). // Cladistic approach to the classification of the genus *Hirrius* Bolívar, 1887 (Orthoptera, Tetrigidae). *Bachelor thesis, Faculty of Science, Zagreb*. [in Croatian]

Deranja, M. (2018) Obrasci rasprostranjenosti širokonosnih trnovratki (Tetrigidae: Cladonotinae) otkrivaju polifiliju potporodice. // Patterns of distribution of groundhoppers (Tetrigidae: Cladonotinae) reveal polyphyly of subfamily. *Bachelor thesis, Faculty of Science, Zagreb*. [in Croatian]

Adžić, K. (2018) Povijesni pregled istraživanja širokonosnih trnovratki (Tetrigidae:Cladonotinae) – od Bolívara do Tumbrincka. // Historical review of (Tetrigidae: Cladonotinae) – from Bolívar to Tumbrinck. *Bachelor thesis, Faculty of Science, Zagreb*. [in Croatian]