Serial endosymbiosis or singular event at the origin of eukaryotes?

Nick Lane

Department of Genetics, Evolution and Environment, University College London

Darwin Building, Gower Street, London WC1E 6BT, United Kingdom

nick.lane@ucl.ac.uk

Abstract

'On the Origin of Mitosing Cells' heralded a new way of seeing cellular evolution, with symbiosis at its heart. Lynn Margulis (then Sagan) marshalled an impressive array of evidence for endosymbiosis, from cell biology to atmospheric chemistry and Earth history. Despite her emphasis on symbiosis, she saw plenty of evidence for gradualism in eukaryotic evolution, with multiple origins of mitosis and sex, repeated acquisitions of plastids, and putative evolutionary intermediates throughout the microbial world. Later on, Margulis maintained her view of multiple endosymbioses giving rise to other organelles such as hydrogenosomes, in keeping with the polyphyletic assumptions of the serial endosymbiosis theory. She stood at the threshold of the phylogenetic era, and anticipated its potential. Yet while predicting that the nucleotide sequences of genes would enable a detailed reconstruction of eukaryotic evolution, Margulis did not, and could not, imagine the radically different story that would eventually emerge from comparative genomics. The last eukaryotic common ancestor now seems to have been essentially a modern eukaryotic cell that had already evolved mitosis, meiotic sex, organelles and endomembrane systems. The long search for missing evolutionary intermediates has failed to turn up a single example, and those discussed by Margulis turn out to have evolved reductively from more complex ancestors. Strikingly, Margulis argued that all eukaryotes had mitochondria in her 1967 paper (a conclusion that she later disavowed). But she developed her ideas in the context of atmospheric oxygen and aerobic respiration, neither of which is consistent with more recent geological and phylogenetic findings. Instead, a modern synthesis of genomics and bioenergetics points to the endosymbiotic restructuring of eukaryotic genomes in relation to bioenergetic membranes as the singular event that permitted the evolution of morphological complexity.

1. The landscape of endosymbiosis in 1967

There can be no doubt that Lynn Margulis's 1967 paper 'On the Origin of Mitosing Cells' (Sagan, 1967) was a seminal, punctuating statement in a century of biology. Little that she wrote was actually new, in that many of the ideas she outlined reached back much earlier in the century. Indeed, reading the paper today, one is struck by how much her cell biology was indebted to the detailed findings of the great cell biologists of the early 20th century, notably Edmund Beecher Wilson (Wilson, 1925) and Clifford Dobbell (Dobell, 1919), as well as Ivan Wallin on the endosymbiotic origin of mitochondria (Wallin, 1927). Wilson, of course, had written a famously withering put-down of early work on endosymbiosis (Wilson, 1925 p. 739); I couldn't help wondering whether Margulis cited him so often deliberately, ironically using his own cell biology to build a compelling contrary case for endosymbiosis. If by 1967 polite biological society was not yet ready to embrace the centrality of endosymbiosis to eukaryotic evolution, after Margulis's paper serious biologists could no longer afford to ignore it. While many aspects of her paper have been debated or contradicted over the ensuing half century, the explanatory power of her main thesis still hits the reader with real force today. And in some respects, Margulis's argument in 1967 was closer to the modern view than her later modifications. Having said that, as this volume will attest, the 'modern view' is by no means unified and uncontested, even if few would any longer support Margulis's case that both mitosis and motility arose from the endosymbiotic acquisition of spirochaetes bacteria (Sagan, 1967).

1.1. Phylogenetic and geological context

Perhaps the most striking and important aspect of her paper was its orchestration of multiple lines of evidence from very different disciplines. Margulis went beyond her own expertise in cell biology to discuss the latest evidence from earth sciences, atmospheric chemistry and genetics, and pointed to the possibilities of phylogenetics. Though written a decade before Carl Woese's revolutionary ribosomal RNA phylogenies were published (Woese and Fox, 1977), she seems to have been aware of (if not citing) Francis Crick's remarks in the late 1950s (Crick, 1958) on the hidden wealth of phenotypic information available from amino acid sequences, and the pioneering work in the early 1960s by Emile Zuckerkandl and Linus Pauling on molecular clocks, which compared the amino acid sequences of haemoglobin chains from different mammals (Zuckerkandl and Pauling, 1965). Margulis writes, for example: "in determining the relationship of two microbes – that is, the amount of time elapsed since they diverged from a common ancestor – we may ask: how many homologous base pair sequences in DNA do they share? The number of mutational steps which occurred to produce one from the other is related to the number of generations elapsed since the two

populations diverged" (Sagan, 1967 p. 249). On the other hand, her estimates of the number of genes and amino acid changes required were startlingly inaccurate. She suggested that the chloroplasts in *Euglena* have "at least 15 different kinds of enzymes" with each one containing about 100 amino acid residues (Sagan, 1967 p. 250, footnote). The chloroplast proteome actually contains as many as 3000 proteins (Qiu et al., 2013), often assembled into giant enzyme complexes, each containing thousands of amino acid residues (Zouni et al., 2001). I find it fascinating the degree to which Margulis and her contemporaries underestimated the molecular complexity of the microbial world, and the multi-subunit protein machines that make it up. This is not a criticism of Margulis, merely a reflection of how much more we know now about protein structures.

But beyond signalling her awareness of the potential, it was too early for phylogenetics to impinge on Margulis's thinking, and later on she distrusted or even rejected the gene-centred view. In 2006, for example, she wrote: "Especially dogmatic are those molecular modelers of the 'tree of life' who, ignorant of alternative topologies (such as webs), don't study ancestors. Victims of a Whiteheadian 'fallacy of misplaced concreteness,' they correlate computer code with names given by 'authorities' to organisms they never see!" (Margulis, 2006). While there may be more than a grain of truth in this, her repudiation of phylogenetics was equally dogmatic, and in stark contrast to her early vision of its possibilities. The fact was that the phylogenetic tree did not correspond well with Margulis's conception of the microbial world, so she preferred to dismiss it altogether in favour of the 'god in the details' of cell biology. Where these two worlds meet, rather than collide, remains a knotty problem which I will explore later.

In contrast, Margulis was arguably decades ahead of her time in considering the detailed geological context of eukaryotic evolution. Preston Cloud, whom she cites extensively, was then reinterpreting the geological record to trace the composition of the atmosphere and oceans from the oxidation of iron and other metals in sedimentary rocks, in relation to fossils of early life (Cloud, 1965). Margulis accordingly split Earth history into a prolonged primordial anaerobic phase, during which oxygenic photosynthesis arose in cyanobacteria (ending in the Great Oxidation Event around 2.4 billion years ago), followed by a long oxygenated phase, during which eukaryotes arose through a succession of endosymbioses. In the 1967 paper, Margulis had the first of these endosymbioses taking place in this oxygenated environment between an unspecified heterotrophic anaerobe and an aerobic bacterial endosymbiont that eventually became integrated as mitochondria in all eukaryotes. Margulis accordingly argued that eukaryotes are fundamentally aerobic, developing their tolerance of oxygen early on through the acquisition of mitochondria (Sagan, 1967).

She was explicit about the basis of the symbiosis, as well as the roles of the two partners involved: "The anaerobic breakdown of glucose to pyruvate along the Embden-Meyerhof pathway occurred in the soluble cytoplasm under the direction of the host genome. Further oxidation of glucose using molecular oxygen via the Krebs cycle… occurred only in the symbiotic mitochondrion under the direction of its own genes" (Sagan, 1967 p. 229). Margulis did not anticipate the level of integration that actually occurs, and seems to have assumed that the mitochondria retained a fully functional genome of their own (capable of controlling replication), as did the host cell. The idea that many mitochondrial genes would eventually be transferred to the nucleus, and that the great respiratory complexes would be composed of proteins encoded by both host and endosymbiont genomes was not easy to predict. Nor was it consistent with an old (and hopeful) prediction that Margulis shared (Wallin, 1927), that mitochondria could be cultured: "If these organelles did indeed originate as free-living microbes, our advancing technology should eventually allow us to supply all growth factors requisite for *in vivo* replication… the *coup de grace* to genetic autonomy" (Sagan, 1967 p. 270). We now know that those 'growth factors' would need to include the protein products of 1500 genes that are located physically in the nucleus (Vafai and Mootha, 2012).

1.2. Oxygen, UV radiation and extinction

Margulis displayed both an unusual breadth of thinking and a curious blind spot in her discussion of atmospheric chemistry. I can only imagine what stimulating conversations she and her cosmologist husband Carl Sagan must have enjoyed over dinner; but it was certainly unusual for biologists to take such a cosmic view of life. Her discussion of prebiotic chemistry is reminiscent of the Miller-Urey experiment (Miller, 1953) in that she called upon a reducing atmosphere containing hydrogen and methane (but trace CO_2); and in some respects she is strikingly modern, invoking cyanide and UV radiation as substrate and driving force. I am not persuaded by the concept of a cyanosulphidic protometabolism driven by UV radiation (Patel et al., 2015), but others do find this approach to the origins of life appealing. Margulis's details, however, lack credibility. She talks about ultraviolet radiation in the upper atmosphere, for example, somehow conjuring ATP (and nucleotides) into existence. In her Table 1, she even refers to 'precellular replicating polynucleotides'. What exactly she had in mind is not clear, but this is close to an RNA world in conception, a hypothesis that was first raised around the same time by Carl Woese (Woese, 1967) and Francis Crick (Crick, 1968). Plainly the idea was in the air. I am struck by how much of the 1967 paper was in harmony with the newest thinking at the time; while Margulis was laying out a radical conception in cell evolution, her

thinking clearly resonated with other leading pioneers of the time. That was not always true later in her life.

At the same time, Margulis seemed oblivious of the link between radiation and oxygen toxicity, first pointed out by Rebeca Gerschman in an emblematic *Science* paper, 'Oxygen poisoning and X-irradiation: a mechanism in common?' (Gerschman et al., 1954). Gerschman's central point was that radiation (including UV radiation) can split water to generate reactive oxygen free radicals, which damage organic molecules including DNA, RNA and proteins. Ground-state oxygen is not particularly reactive or toxic, despite being a free radical itself, as it can only accept single electrons from relatively willing donors, such as Fe^{2+}. On accepting single electrons, the same reactive oxygen species are formed that are produced by irradiation of water – superoxide ($O_2^{\bullet-}$), hydrogen peroxide (H_2O_2) and the hydroxyl radical ($OH^{\bullet}$). Only the hydroxyl radical is aggressively reactive; and that is more likely to be formed directly by a single-electron oxidation of water than the three-electron reduction of oxygen (Lane, 2002). So it is ironic that Margulis credits UV radiation as the driving force behind prebiotic chemistry, and yet considered oxygen to be "lethal to early self-replicating systems" (Sagan, 1967 p.258).

Over evolutionary time, Margulis plainly saw oxygen as a kind of a binary geological switch, whereby global conditions were either anoxic or aerobic (with limited anaerobic refugia), leading to what she later termed the oxygen 'holocaust', in which most 'primitive' anaerobes fell extinct (Margulis and Sagan, 1997). In fact, there is no evidence whatsoever for an oxygen holocaust (Lane, 2002, 2011). Quite the opposite. By oxidising minerals eroded from terrestrial rocks, rising atmospheric oxygen levels produced a much greater flux of alternative electron acceptors, such as sulphate and nitrite (Canfield, 1998; Knoll et al., 2016). The possibilities for anaerobic lifestyles multiplied (Mentel and Martin, 2008). Biochemical and geochemical feedbacks (for example, the ecological expansion of sulphate-reducing bacteria) meant that the oceans remained largely anoxic (either sulphidic or ferruginous, particularly below the photic zone) for more than a billion years after the Great Oxidation Event (Boyle et al., 2013). The absence of a binary switch from anaerobic to aerobic conditions, combined with the later wholly unanticipated findings from comparative genomics (reviewed in Williams et al, 2013), which Margulis never accepted, means that there is a conflict between two distinct conceptions of eukaryotic evolution. In essence, the alternative views contrast a polyphyletic and gradualistic framework for evolution, driven by microbial collaborations following an environmental bottleneck (rising oxygen), against a singular and improbable origin of eukaryotes, stemming from a restrictive bottleneck – the physical structure of prokaryotic cells (**Fig. 1**) (Lane,

2005; Lane and Martin 2010; Lane 2015). The predictions that emerge from each view are radically different, despite endosymbiosis playing a central role in both. I will explore how well each hypothesis corresponds to modern data, drawing, as exemplified by Margulis, on cell biology, phylogeny and earth history.

2. The problem with oxygen as an environmental bottleneck

I will focus here on Margulis's central argument that the combination of oxygen and endosymbiotic cooperation together drove the evolution of the eukaryotic cell, as the 1967 paper contains a set of implicit predictions that resonate with her whole approach to cell evolution later in life. By 'implicit predictions', I am not referring to the explicit predictions laid out in the paper, many of which have been verified. For example, Margulis notes that if her theory is correct then all eukaryotic cells must be seen as multi-genomed systems (Sagan, 1967 p. 271). This is not strictly true (in that hydrogenosomes and mitosomes have typically lost all their genes) but in terms of eukaryotic origins it is certainly true that all known eukaryotes either possess, or once had and later lost, mitochondria (Keeling, 1998; Embley and Martin, 2006; Müller et al., 2012; Archibald, 2015). It is therefore correct to view eukaryotes as ancestrally multi-genomed systems. Margulis also predicted that free-living relatives of endosymbiotic bacteria would be found among modern bacterial groups, and pointed to aerobic cytochrome-containing bacteria as relatives of mitochondria, blue-green algae (cyanobacteria) as free-living relatives of chloroplasts (albeit she argued there were repeated acquisitions of plastids from prokaryotes) and spirochaetes as the ancestors of eukaryotic flagella. With the exception of spirochaetes (and the single primary origin of plastids) her predictions were broadly correct, and almost universally accepted today. Her fame is founded on these ideas, which are of undisputed significance; but there are some difficulties with the assumptions that lie beneath them.

2.1 Implicit predictions from serial endosymbiosis

Most illuminatingly, Margulis also predicted that some searches would continue to be futile, that certain missing links would never be found. Amongst these, she numbered organisms containing chloroplasts but no mitochondria (which have still never been found), eumitotic organisms with bacterial flagella (also absent), eumitotic fossils from anaerobic times (which is not really true, as eukaryotes arose in dysoxic conditions) and most notably, eumitosis in all eukaryotes. She clarifies this remark by stating that if it is actually found, 'eumitosis' will clearly be analogous, rather than homologous, to eumitosis in higher eukaryotes, and she compares this possibility with reports of sexuality in the dinoflagellate *Noctiluca*. There is an implicit assumption underlying this prediction,

which structures Margulis's entire conception of eukaryotic evolution: it is that the succession of endosymbioses, and the ensuing gradualistic evolution that depended on these endosymbioses (such as the evolution of mitosis following the acquisition of spirochaetes) gave rise to numerous evolutionary intermediates that can still be found among the rich tapestry of eukaryotic protists today. This thinking permeated Margulis's writing many years later, for example in her assertions that hydrogenosomes do not derive from mitochondria (Margulis, 2005), but rather were separate acquisitions, perhaps deriving from *Clostridia* as had originally been proposed by Miklos Müller (Lindmark and Müller, 1973).

In short, the implicit prediction underlying the serial endosymbiosis hypothesis is that there should be numerous evolutionary intermediates that never acquired particular endosymbionts. This view remained constant even when specific details changed dramatically over decades. For example, in the 1967 paper, Margulis explicitly predicted that the premitotic eukaryotes represented a period before the acquisition of spirochaetes. She rejected the possibility that the 'aberrant' forms of cell division found in *Amoeba*, *Euglena* and *Tetramitus* were degenerate, deriving from ancestors that had already evolved eumitosis, on the basis that these groups also seemed to be primitively asexual: "They are probably not degenerate phytoflagellates, but eukaryotic organisms which are premitotic in the sense that they branched off the main lines of higher cell evolution before eumitosis evolved." (Sagan, 1967 p.233). She applied similar arguments to the acquisition of plastids. "The diversity of cell structure and the life cycle in lower eukaryotic algae imply that different photosynthetic prokaryotes (proto-plastids) were ingested at various times during the evolution of eumitosis." (Sagan, 1967 p.247). But some years later she reversed her argument, in 2005 claiming that spirochaetes were acquired before mitochondria, and that some ciliates were primitively amitochondriate (Margulis et al., 2005). Ironically, her argument echoed the 'archezoan' hypothesis of Tom Cavalier-Smith, which had held that at least some primitive eukaryotes had never possessed mitochondria (Cavalier-Smith, 1987, 1989). By 2005, two decades of work had established that all putative archeozoa had organelles derived from mitochondria (hydrogenosomes or mitosomes), hence had evolved by reductive evolution from more complex ancestors (Keeling, 1998; Embley and Martin, 2006; Archibald, 2015). By 2005 even Cavalier-Smith had abandoned the archezoa hypothesis (Cavalier-Smith, 2002); yet by then Margulis was arguing that "descendants of these amitochondriate cells (archaeprotists) today thrive in organic-rich anoxic habitats where they are amenable to study" (Margulis et al., 2005). They have been studied; and another decade later, there is still no evidence for primitively amitochondriate eukaryotes, even if some protists (notably *Monocercomonoides*) have entirely lost the organelle (Karnkowska et al., 2016).

2.2 Environmental bottlenecks and polyphyletic origins

So why did Margulis switch from arguing, in 1967, that no primitively amitochondriate eukaryotes existed, to positing in 2005 that they were abundant yet unexplored? It was not merely a matter of data, as the data showed the opposite. The answer might lie in her conception of Gaia, the idea she developed with James Lovelock, that the living world is composed of a network of microbial communities (Lovelock and Margulis, 1974). Evolution was not about selfish genes and aggressive male competition, but softer, more feminine virtues: "Life did not take over the globe by combat, but by networking: by cooperation, interaction, and mutual dependence between living organisms." (Margulis and Sagan, 2002). But if eukaryotes arose through a tapestry of symbioses, then there ought to have been different symbioses in diverse environments. Different collaborations. Different intermediates. A richness to cell evolution expressed through a myriad of distinct, independent cell structures. I suspect that is why Margulis always saw intermediates, and equally why she contested the troubling conclusions of phylogenetics – that there aren't any intermediates (Keeling, 1998; Embley and Martin, 2006; Archibald, 2015). Certainly, there are many diverse endosymbioses between eukaryotes and prokaryotes in different environments, but all of them are between 'fully-fledged' eukaryotic cells (which either have or once had mitochondria, endoplasmic reticulum, nucleus, mitosis, meiosis, etc) and bacteria (McCutcheon and Moran, 2012, Wernegreen, 2012, Archibald, 2015). If eukaryotes *arose* by networking, then the prediction is that different types of complex 'eukaryotic' cells ought to evolve via different collaborations in distinct environments (**Fig 1 a, b**). In other words, eukaryotes should have polyphyletic origins. In arguing that some eukaryotes had never acquired mitochondria (Margulis et al., 2005) or had never evolved eumitosis through an endosymbiosis with spirochaetes (Sagan, 1967) that was precisely Margulis's point.

This emphasis on networking also explains why oxygen was central to her original conception of eukaryotic evolution: oxygen was an *environmental* gatekeeper. There is no doubt that bacteria and archaea do collaborate, through rich syntrophic relationships. That does not mean they do not compete. I think Margulis was simply wrong, later in life, to claim that "symbiosis has nothing to do with cost or benefit. The benefit/cost people have perverted the science with invidious economic analogies" (Margulis, 2006). Few serious biologists could accept that statement. But putting that aside, and thinking in Margulis's own terms – if evolution is made up of collaborating prokaryotes, which in turn form the chimeric eukaryotic cell, then why do we find clear evidence of prokaryotes in the geological record dating back nearly 4 billion years (Knoll et al., 2016), but no trace of eukaryotes before 1.5–2 billion years ago (Knoll et al., 2006)? Margulis was equally clear about this in 1967,

giving a date for the evolution of eukaryotes, based on the fossil evidence at that time, of about 1.2 billion years ago; and the origin of life before 3.1 billion years ago, again a gap of two billion years. If prokaryotic evolution was fundamentally about collaboration, then Margulis needed to explain this long delay before they gave rise to eukaryotic cells. In 1967, she saw the explanation, then perfectly reasonably, in terms of an environmental bottleneck: the cellular complexity of eukaryotes was not possible in the absence of oxygen. She assumed that it took the cyanobacteria aeons to transform the planet into a home fit for eukaryotes, eventually via an oxygen catastrophe, hence the delay. The mitochondrial-host endosymbiosis, Margulis suggested with great prescience (before the discovery of the archaea), "might have resulted in the typical eukaryotic phospholipid membrane and steroid synthesis, and in particular, the formation of a nuclear membrane and endoplasmic reticulum." (Sagan, 1967 p. 229). She went on to postulate that "the greater amounts of energy available after the incorporation of the mitochondrion resulted in large cells with amoeboid and cyclotic movement. However, the diversity in types and amounts of proteins such cells could make would have been limited by the amount of DNA available to administer protein synthesis." (Sagan, 1967 p.229). So mitochondria provided more energy, enabling cells to become larger and eventually support more genes and proteins.

Margulis here put in a nutshell some of the arguments that Bill Martin and I advanced in 2010 (Lane and Martin, 2010), which I will outline in the final section of this article, with one crucial exception. Margulis saw the boost in energy availability, and ultimately capacity for protein synthesis, not in the structural reorganization that gave rise to the eukaryotic cell (and specifically the topology of membranes in relation to genes) but in the capacity to respire oxygen. That is not true, and cannot be true, for many complex eukaryotes have mitochondria but do not respire aerobically, whereas many bacteria and archaea are aerobic, yet they never become large and morphologically complex (Martin et al., 2001; Lane and Martin, 2010; Müller et al., 2012). Margulis was tantalisingly close to the answer, but the problem with oxygen, as with any other environmental gatekeeper, is the implicit prediction that there should be polyphyletic origins of morphological complexity. While it is true, as Margulis contended, that cyanobacteria evolved in an anaerobic world, they still thrive in aerobic conditions today. Yet the largest genomes in cyanobacteria are about 12 Mb, whereas eukaryotic algae have genome sizes ranging up to 150,000 Mb, many orders of magnitude greater (Elliot and Gregory, 2015), despite an equivalent phototrophic lifestyle. There is no good reason why mitochondria should be needed to adapt to oxygen – the FeS clusters in respiratory chains are in fact the worst generators of reactive oxygen species (Murphy, 2009). And if oxygen is the key, there is no good reason why there should not have been multiple origins of complex cells, entailing many

different symbiotic combinations of prokaryotes adapted to diverse ways of life, from phototrophy to osmotrophy to predation (**Fig. 1a, b**). One would expect that algae should arise from cyanobacteria, fungi from osmotrophic bacteria and phagocytes from predatory bacteria. Yet that is not what happened. The phylogenetics that Margulis ultimately rejected show unequivocally that the common ancestor of eukaryotes was tantamount to a modern cell with mitochondria, a dynamic cytoskeleton, nucleus and endoplasmic reticulum, mitosis and meiosis; more or less all eukaryotic traits (Koonin, 2010; 2015). That was not what Margulis had ever predicted, and is not consistent with the serial endosymbiosis hypothesis. So what did happen?

3. A restrictive bottleneck: restructuring genomes in relation to bioenergetic membranes

Two major factors have emerged over the last decades, which have together turned the landscape of endosymbiosis as conceived by Margulis in 1967 on its head. These factors do not in themselves provide an answer, but they structure the question differently.

3.1 The modern context for endosymbiosis

First, it is now clear from phylogenetics that the eukaryotes are a derived domain: the outcome of an endosymbiosis between an archaeal host cell and a bacterial endosymbiont (Williams et al., 2013). The bacterial symbiont was probably related to the proteobacteria, but does not correspond exactly to any modern group (Müller et al., 2012). Recent research suggests that the host cell was most likely related to the Lokiarchaeota (Spang et al., 2015, Zaremba-Niedzwiedzka, 2017) but whether this was a relatively complex archaeon, capable of some form of rudimentary phagocytosis (Martijn and Ettema, 2013), or a common-or-garden variety, for example a hydrogen-dependent autotroph (Sousa et al., 2016), is still disputed. The problem is that all the eukaryotic supergroups share virtually all the same genes specifying eukaryotic architecture and behaviour (such as nuclear pore complexes, or a two-step meiosis). So there is a phylogenetic 'event-horizon' between the last eukaryotic common ancestor, which had everything, and all prokaryotes, which in these terms have next to nothing (Lane, 2011, 2015). This perplexing fact needs an explanation: why should eukaryotes share such great complexity, which is not obviously related to a particular environment or lifestyle, yet bacteria and archaea show little if any propensity to evolve any of those traits? The so-called archezoa, once proposed as possible evolutionary intermediates, turn out to be derived from more complex ancestors (Keeling, 1998; Embley and Martin, 2006; Archibald, 2015); but their abundance demonstrates that simple eukaryotes are not necessarily outcompeted to extinction by more sophisticated cells (Lane, 2011). We now know that almost all variation between eukaryotes

reflects secondary adaptations, and virtually none reflects steps along the path of eukaryogenesis. The absence of evolutionary intermediates therefore seems to be telling us something important – perhaps that the early ancestors of eukaryotes comprised a small, unstable, rapidly evolving sexual population (Lane, 2011). That is what one would predict if there were indeed a bottleneck at eukaryotic origins; but very specifically a *restrictive* bottleneck, a permissive change in cell structure, not an environmental bottleneck.

Second, there was no binary switch from anaerobic to aerobic conditions. The eukaryotes apparently arose sometime during the 'boring billion', the period after the Great Oxidation Event and before the Neoproterozoic Snowball Earths, when the oceans were largely anoxic, with low levels of oxygen in the photic zone and atmosphere (Knoll et al., 2006). So the eukaryotes probably evolved in dysoxic conditions, and this is reflected in the metabolism of many anaerobic or facultatively aerobic eukaryotes (Mentel and Martin, 2008; Müller et al, 2012). The fact that hydrogenosomes and mitosomes almost invariably retain the same small subset of enzymes implies that they inherited them from a common ancestor (Martin, 1999; Martin et al., 2001), even if the gene trees themselves display some conflict that some have interpreted as lateral gene transfer from bacteria (Nyvltova et al., 2014). While eukaryotes might tend to acquire similar genes in equivalent environments by LGT, there is no strong reason to suppose that eukaryotes should repeatedly acquire an identical subset of genes from bacteria, to the exclusion of the myriad alternative forms of anaerobic respiration in bacteria. Even the subset of bacteria living in hydrothermal systems have more than 150 different electron acceptors (Amend and Shock, 2001) compared with just a handful in all known eukaryotes (Martin et al., 2001). It makes more sense to view the dissonance in gene trees as an artefact, or divergent rates of fixation in different environments, or convergent evolution at the level of gene sequences (for example with different groups of protists adapting to similar conditions) (Martin, 1999). The simplest explanation for the fact that the entire eukaryotic domain has the same metabolic versatility as a single facultatively anaerobic bacterium is that eukaryotes actually did acquire their metabolic versatility from a single bacterial endosymbiont living in a dysoxic environment (Martin 1999; 2001). If so, the origin of eukaryotes had little if anything to do with oxygen, and much to do with the acquisition of mitochondria (Martin, 1999). This conclusion is supported by the fact that at a cellular level there is virtually no difference in complexity between aerobic and anaerobic eukaryotes (Müller et al., 2012). As noted above, the acquisition of mitochondria produced a permissive change in cell structure, which freed eukaryotes from the constraints acting on all prokaryotes. A restrictive bottleneck pleasingly explains the monophyletic origin of eukaryotes.

3.2 Endosymbiosis between prokaryotes as a restrictive bottleneck

The idea of a restrictive bottleneck at the origin of eukaryotes is not new. Cavalier-Smith's proposal that the catastrophic loss of the cell wall drove the 'quantum evolution' of archaea and eukaryotes corresponds to a rare change in cell structure, a restrictive bottleneck (Cavalier-Smith, 2002). So does the hypothesis that prokaryotic replication is limited by a single circular chromosome attached to the cell membrane; having straight chromosomes with multiple origins of replication solves this topological problem via a rare change in cell structure (Maynard Smith and Szathmary, 1995). The issue in both these cases is that there are plenty of examples of prokaryotes lacking cell walls (Errington, 2013), or with straight chromosomes and multiple origins of replication (Barry and Bell, 2006), yet in no case did they evolve eukaryotic complexity. The evolution of phagotrophy in primitive eukaryotes has also been suggested as a restrictive bottleneck (Cavalier-Smith, 2002; Cavalier-Smith, 2014), but the evidence is equivocal. Phagocytosis might in fact have arisen independently in three modern groups of eukaryotes, after the last eukaryotic common ancestor (Yutin et al, 2009). An early evolution of phagocytosis is not consistent with proteomic evidence either. If the nucleus evolved in response to shear stresses from phagocytosis, for example, and mitochondria were acquired later by a phagocytic host cell, then the nucleus should be a host cell innovation (Cavalier-Smith, 2002). But in fact the nuclear proteome is chimeric: some nucleolar, laminar and nuclear-pore proteins derive from archaea and others from bacteria (Staub et al, 2004; Mans et al, 2004, McInerney et al, 2011), implying that these nuclear structures first evolved in the context of an endosymbiosis between an archaeon and a bacterium.

All the evidence discussed above is consistent with a different type of restrictive bottleneck at the origin of eukaryotes, however – a singular endosymbiosis between two prokaryotes, in which an archaeal host cell acquired a facultatively anaerobic bacterial symbiont, as for example posited in the hydrogen hypothesis (Martin and Müller, 1998). The idea has great explanatory power. If true, essentially all eukaryotic complexity from the nucleus to sex and phagocytosis (none of which is known in bacteria or archaea) evolved in the context of an endosymbiosis between prokaryotes (Martin and Koonin, 2006; Lane, 2005, 2011). This can explain why, for example, a plant cell and an animal cell should have the same basic cellular architecture and behaviour despite utterly different lifestyles. Adaptation was not to some shared external environment or lifestyle, but rather to the *internal* environment: the presence of bacterial endosymbionts in an archaeal host cell.

3.3 Were mitochondria acquired early?

For eukaryotic evolution to have been driven by the acquisition of mitochondria, mitochondria must obviously have been acquired at the very beginning of eukaryotic evolution. This is wholly consistent with the phylogenomic evidence that the host cell was an archaeon (Williams et al., 2013; Spang et al., 2015; Zaremba-Niedzwiedzka et al., 2017): an endosymbiosis between two prokaryotes was the *sine qua non* for the evolution of all eukaryotic traits. But this conception is not consistent with the recent suggestion that mitochondria were acquired later in eukaryotic evolution, based on the (disputed) observation that the genes involved in mitochondrial respiration are more closely related to bacteria than other eukaryotic genes that have bacterial homology, such as those expressed in the nucleus and endoplasmic reticulum (Pittis and Gabaldon, 2016). The statistical basis of these findings has been challenged (Martin et al., 2016), but even if correct, the evolutionary distance says nothing about the time of acquisition. If the genes involved in oxidative phosphorylation are more similar to their bacterial relatives, then that most likely reflects the strength of purifying selection, which maintains exactly the same function (oxidative phosphorylation) in an equivalent setting – a membrane with bacterial lipids and strong electrical potential. By eliminating most variants, strong purifying selection minimizes genetic distance. In contrast, genes deriving from the same bacteria, but now expressed in the nuclear membrane or endoplasmic reticulum, must have been subject to strong adaptive selection for an entirely new function in a completely novel setting, as the ER and nucleus, along with their proteome, does not exist in bacteria or archaea. Similar reasoning applies to ribosomes, which are larger and more complex in eukaryotes than prokaryotes (Dinman, 2008), and again operate in a new compartment, the eukaryotic cytosol, frequently in association with novel endomembrane systems, such as the rough ER. Adaptive selection necessarily increases evolutionary distance. So even if it is true that the genetic distance between bacterial genes and their eukaryotic homologues is greater for genes that are not expressed in the mitochondria, this does not imply that mitochondria were acquired later in eukaryotic evolution. I should note, too, that there could have been strong purifying selection over the ensuing 1.5 billion years, making any evolutionary signal difficult to detect.

Nor does a singular origin of eukaryotes preclude other endosymbioses occurring at a later stage, as undoubtedly happened. Plastids, for example, derive from cyanobacteria, as Margulis knew (Sagan, 1967). She was wrong, as it happens, in that there was just one primary endosymbiosis in which cyanobacteria were engulfed, followed by secondary and tertiary endosymbioses in various groups, in which phagocytic protists engulfed eukaryotic algae (Archibald, 2009). There are also hundreds of examples of bacterial endosymbionts and parasites living in eukaryotic host cells (McCutcheon and Moran, 2012, Wernegreen, 2012). As noted earlier, all these endosymbioses are with 'fully fledged'

eukaryotes, which had already evolved the nucleus, complex endomembrane systems, dynamic cytoskeleton, mitosis, meiotic sex and much more (McCutcheon and Moran, 2012, Archibald, 2015). They did not contribute to the evolution of these basal eukaryotic traits. So what did contribute to the evolution of these traits? Only the mitochondria unequivocally derive from an endosymbiosis that took place before the evolution of last eukaryotic common ancestor (Embley and Martin, 2006; Archibald, 2015). Assuming that mitochondria were acquired by an archaeal host cell – something similar to Lokiarchaeota, as phylogenetics suggest (Spang et al., 2015, Zaremba-Niedzwiedzka, 2017) – then the host cell did not have a nucleus or any of these basal eukaryotic traits (Sousa et al, 2016). The acquisition of mitochondria must have shifted the balance of selection pressures acting on cells, in part by increasing energy availability (discussed below), but also by obliging the host cells and their endosymbionts to align life cycles and resolve conflicts (Blackstone, 2013) through the evolution of traits such as the nucleus (Martin and Koonin, 2006), sex (Lane, 2011, Radzvilavicius and Blackstone, 2016) and mating types (Hadjivasiliou et al, 2012; 2013). The fact that eukaryotes share so many basal traits, none of which are found in comparable form in bacteria or archaea, suggests that they evolved in a small, fast-evolving, proto-sexual population (Lane, 2011; Lane, 2015). If the population was large and stable, it should have become structured in space, leading to divergence and speciation (Lane, 2011; Lane 2015). Likewise, if eukaryotes evolved over tens or hundreds of millions of years (as opposed to perhaps a few million years), one would expect to see at least some surviving evolutionary intermediates (Lane, 2011; Lane 2015). But there are none (Keeling, 1998; Embley and Martin, 2006; Archibald, 2015). From a phylogenetic point of view, the discontinuity between prokaryotes and eukaryotes looks like an unbranching trunk (**Fig. 2c**). The critical point is that the starting point was radically different: rather than natural selection operating on populations of individual bacteria or archaea, it was acting on populations of prokaryotes with endosymbionts. The dominant selective forces driving eukaryotic evolution arose from within the cell, not the external environment (Lane, 2015).

3.3 The singular origin of eukaryotes

All of this points to the singularity of eukaryotic origins. The acquisition of mitochondria was not equivalent to the acquisition of plastids or any other endosymbionts, as the host cell that acquired mitochondria was not a fully-fledged eukaryotic phagocyte but an archaeon that was unlikely to be capable of phagocytosis at all (Sousa et al 2016). That of course begs the thorny age-old question, how did the mitochondria physically get inside a non-phagocytic host cell? The answer is, we don't know. But we do know that it is possible, for there is one known example of free-living bacteria – with cell walls – that have bacterial endosymbionts (Wujek, 1978). There are other, more equivocal,

examples of prokaryotes with bacterial endosymbionts (von Dohlen et al., 2001; McCutcheon and von Dohlen, 2012), and bacterial endosymbionts in fungi (Minerdi et al., 2002), which are no more phagocytic than bacteria. So we don't know how they got in, but we do know that it is possible, if rare. And this explanation sits far more comfortably with the peculiar trajectory of eukaryotic evolution. A rare event can in principle explain why all eukaryotic cells share a complex common ancestor that only arose once; why all eukaryotes are monophyletic rather than arising through many different endosymbioses; why there are no known surviving evolutionary intermediates; and why neither bacteria nor archaea show any tendency to evolve those complex morphological and behavioural traits that are shared by all eukaryotes. An endosymbiosis between prokaryotes could explain not only the singularity of eukaryotic origins, but also, potentially, many of these complex eukaryotic traits (Martin and Koonin, 2006; Lane, 2005, 2015). All of that stems from the game-changing endosymbiosis at the origin of eukaryotes, which set in motion a different set of selective driving forces leading to the gradualistic evolution of eukaryotic traits (albeit over millions of years rather than tens or hundreds of millions of years). The later acquisition of plastids, and multiple other bacterial and algal endosymbionts merely adds finesse to the story. Later endosymbioses did not fundamentally alter the structure of cells in the way that the mitochondrial endosymbiosis did (**Fig. 1c**).

4. Why membrane bioenergetics holds the key to complexity

Mitochondria gave eukaryotes several *orders of magnitude* more energy per gene (Lane and Martin, 2010). This massive rise in energy availability had nothing to do with oxygen and very little to do with increasing the internal surface area of bioenergetic membranes (Lane, 2011; 2014; 2015; Lane and Martin, 2016), although this has caused some confusion; Lynch and Marinov 2015; 2017). Bacteria often have complex internal membranes such as the thylakoid membranes in cyanobacteria, yet they do not become large and complex, certainly not on the scale of eukaryotes. The key point is that endosymbiosis allowed a restructuring of genomes in relation to bioenergetic membranes (Lane and Martin, 2010). In effect, eukaryotes have multi-bacterial power without the genomic overheads. Mitochondria lost the great majority of their genes, ultimately retaining only a handful of genes that have never been lost without cells also losing the ability to perform oxidative phosphorylation (Allen, 1993; Allen, 2003; Allen 2017). The reason that mitochondria (and chloroplasts) never lost these genes probably relates to a strict requirement to control membrane potential locally and rapidly, as postulated in John Allen's CORR (co-location for redox regulation) hypothesis (Allen, 1993; Allen, 2003; Allen 2017). A strict requirement for genes to control membrane potential is borne out by a few examples of giant bacteria, notably *Epulopiscium* and

Thiomargarita, which are visible to the naked eye and far larger than most eukaryotic protists (Schulz and Jorgensen, 2001; Schulz, 2006, Mendell et al., 2008). These giant cells invariably display extreme polyploidy, in which tens of thousands of copies of the complete genome are positioned at regular intervals immediately next to the plasma membrane (Mendell, 2008, Angert, 2012). Even large cyanobacteria (which are substantially smaller than such behemoths) have hundreds of copies of their complete genome next to their thylakoid membranes (Griese, 2011; Angert, 2012). The cost of expressing all this repetitive DNA means that giant bacteria have no more energy per haploid copy of each gene than small bacteria such as *E. coli* (Lane and Martin, 2010). And they can only manage that because most of their inner volume is metabolically almost inert – a giant vacuole in the case of *Thiomargarita* (Schulz, 2006) and sporulating daughter cells in the case of *Epulopiscium* (Ward, et al., 2009). There is no comparison to the extreme morphological complexity and energy-guzzling machinery of eukaryotic cells. I stress that *all* the (few) giant bacteria known *always* display extreme polyploidy, with thousands of copies of their fairly small (~3 Mb) bacterial genomes (Angert, 2012). These genomes never vary in gene content, because genomes are not independent, self-replicating entities, but are inert gene banks acted upon by proteins. There is a genomic symmetry to bacteria, in that their polyploid genomes are equal in size and gene content. The scaling of polyploid genomes with prokaryotic cell volume is ultimately what prevents both bacteria and archaea from attaining eukaryotic complexity (but has been ignored by some; Lynch and Marinov, 2015; 2017; Lane and Martin, 2016).

4.1 The implications of genomic asymmetry

Endosymbiosis breaks up this genomic symmetry. Bacteria are more than inert genomes – they are populations of cells, which compete and undergo selection. The fact that mitochondria lost most of their genes means that the costs of protein synthesis shrank, while their ATP output remained high. The energy savings accruing from gene loss in mitochondria equate to the costs no longer incurred by necessary protein synthesis. In a homeostatic intracellular environment, endosymbionts can afford to lose unnecessary traits such as the cell wall, along with the genes and proteins needed to produce them. If each of 100 endosymbionts with typical bacterial genomes of 4000 genes were to lose 5% of their genome (200 genes), the energy savings from not making those proteins would be around 50 billion ATPs (Lane, 2014). Assuming a conservative life cycle of 24 hours, that gives an energy saving of 580,000 ATPs per second! In fact, mitochondria lost much more than 5% of their genes (not including those that were transferred to the nucleus but continued to perform equivalent tasks). And there are far more than 100 mitochondria in eukaryotic cells; the giant *Amoeba proteus* has as many as 300,000 (Daniels and Breyer, 1968). The energy savings are astronomical, and stem

directly from a genomic asymmetry, in which tiny mitochondrial genomes support, energetically, the expansion of the host-cell genome (Lane, 2011). There may be no more DNA in total than in giant bacteria, but its distribution is radically different. I would argue that genomic asymmetry is a better defining feature of eukaryotes than the nucleus alone. As Margulis observed (Sagan, 1967 p. 271), eukaryotes are fundamentally multigenomed cells, and their genomes are not equal in size. This is the key point overlooked by Lynch and Marinov (2015, 2017), who in their estimates of the energetic costs of scaling omitted to discuss either extreme polyploidy in giant bacteria, or extreme polyploidy of mitochondrial DNA in eukaryotes (Lane and Martin 2016).

There was no necessity for all these energy savings to be spent, but the fact is that they were: they were spent on supporting the greater size and morphological complexity of eukaryotes. When Bill Martin and I say that eukaryotes have 100,000 times more energy per gene than bacteria, that does not mean they harbour genome sizes 100,000 times larger (as has been claimed; Booth and Doolittle, 2015; Lane and Martin, 2015), rather that the energy available for gene expression – protein synthesis – is increased by that factor. Eukaryotes could support 100,000 times more protein synthesis from the same number of genes, or the same gene expression from 100,000 more genes, or any combination of the two, as previously pointed out (Lane, 2014; 2015). In fact, eukaryotes typically have around four times as many genes as bacteria, but these are expressed at far higher levels, if only because eukaryotic cells are on average 15,000 times larger. A simple example is the ribosome. A single *E. coli* cell has up to 13,000 ribosomes, while a single liver cell has 13 *million* on the rough endoplasmic reticulum alone – a factor of 1,000-10,000 times more (Lane and Martin, 2010). So 'energy per gene' specifically means the energy availability for protein synthesis and does not imply expansion in genome size by that factor (Lane, 2015; Lane and Martin, 2016). The acquisition of mitochondria meant that eukaryotes were no longer limited in their structure from becoming larger and more complex. They could accumulate new gene families (and there were some 3000 new gene families in the last eukaryotic common ancestor; Koonin et al., 2004; Fritz-Laylin et al., 2010), larger proteins (Brocchieri and Karlin, 2005), and higher gene expression. They could also accumulate far more DNA and regulatory elements (Elliot and Gregory, 2015). That does not mean that eukaryotes necessarily have large flabby genomes – some have plainly been selected for genomic streamlining. It just means that they can – something that we never see in prokaryotes. The largest genome size known in bacteria and archaea is less than 15 Mb in size; the largest eukaryotic genome sizes range up to 150,000 Mb, a 10,000-fold difference (Elliot and Gregory, 2015), which is not consistent with a simple continuum of complexity between prokaryotic and

eukaryotic domains, as claimed by some (Lynch and Marinov, 2016, 2017; see Lane and Martin, 2016).

4.2 How eukaryotes can lose their mitochondria

If that is the difference that mitochondria make, then how did some eukaryotes manage to lose their mitochondria altogether yet remain relatively complex? The answer relates to selection pressures. In bacteria there is little if any benefit to being a little larger, having a little more membrane, and more ATP; those cells tend to lose out to smaller, more streamlined cells that replicate faster (Vellai et al., 1998). Bacteria are not deficient in ATP, so having more provides little benefit (Lane, 2011; 2015). This tendency is probably most extreme among obligate fermenters. Bacterial fermenters compete with other cells that can extract more energy from the same substrate, hence can keep on growing for longer. Obligate fermenters among bacteria are therefore obliged to compete by growing fast and replicating quickly, and so are usually among the smallest and most genomically streamlined cells (Makarova and Koonin, 2007).

But the endosymbiotic origin of the eukaryotic cell had nothing to do with ATP – it could not have done, for no bacteria export ATP to the surroundings. Most probably, to be stable over evolutionary time, the endosymbiosis was based on a metabolic syntrophy, in which the endosymbiont provided the host with the *substrates* needed for growth – for example H_2 gas in the case of the hydrogen hypothesis (Martin and Müller, 1998). In which case, the more endosymbionts, the more substrates, and the more growth. Being larger and evolving transport pathways to deliver organics to bacterial endosymbionts would have had an advantage from the beginning, already breaking the standard selection pressures, because adaptation was in response to *internal*, not external, factors. For the host cell, large size, improved transport networks and greater internal complexity would be favoured generation after generation (Lane, 2015).

Eventually, with the evolution of an ATP/ADP translocator, ATP became an important factor. Now eukaryotes had, if anything, too much ATP, which (to control the membrane potential of their own mitochondria) needed to be rapidly recycled back to ADP (Radzvilavicius and Blackstone, 2015). This excess ATP was spent, most readily, on infrastructure projects such as a dynamic cytoskeleton and membrane trafficking, which did not interfere with flux through metabolic pathways, but which facilitated movement, endocytosis, and ultimately phagocytosis (Garg and Martin, 2016). Phagocytosis must have required the *de novo* evolution and extremely high expression of hundreds if not thousands of new genes – an energy investment that seems to be beyond any prokaryote

(Lane and Martin, 2010; 2016). But once all this machinery had evolved, and the benefits were in place, the energy demands to *maintain* a phagocytic lifestyle were surely transformed. The reason is the nature of phagocytosis itself. So long as phagocytes live in an organic-rich environment they can survive by fermentation alone, because unlike bacteria they do not need to compete in terms of growth rate; they simply eat the opposition. So phagocytes can survive by fermentation alone, losing their mitochondria, and there are examples of morphologically simple phagocytes with modestly large genomes among the archezoa (e.g. *Tritrichomonas foetus* has 177 Mb; Zubacova et al., 2008), all of which derive from more complex ancestors. There is a critical distinction between the *origin* of phagocytosis, which required mitochondria (providing both energy and selective driving force) and the *retention* of phagocytosis, which merely required enough prey to fuel fermentation, allowing simpler phagocytes to dispense with their mitochondria altogether. This distinction is sharp but has sometimes been overlooked (Booth and Doolittle, 2015; Lane and Martin, 2016).

5. Conclusions

Lynn Margulis's 1967 paper was genuinely seminal: it changed the way that biology was understood. Until then, few people had taken endosymbiosis seriously as a driving force in evolution. After that, no serious biologist could ignore it. Margulis was fundamentally correct that the mitochondria and plastids derive from proteobacteria and cyanobacteria respectively, and her wide-ranging arguments were often compelling and made sense of a great deal of cell biology that must have seemed opaque to earlier generations. She was almost certainly wrong in one important aspect of her paper, the acquisition of spirochaetes leading to the evolution of centrosomes, flagellae and mitotic spindles (Carvalho-Santos et al, 2011). Perhaps the strangest aspect of her 1967 paper relates to precisely that: in Table 3, she considers the paternal inheritance of mitochondria as a possible mechanism for ensuring the co-transmission of mitochondria with the flagellum of sperm (which she argued were derived from endosymbiotic spirochaetes). Margulis barely mentions the maternal transmission of mitochondria, which was well known at the time (Gibor and Granick, 1964), but she seems to have allowed her judgement to be led astray by a focus on spirochaetes. But all great scientists are wrong about aspects of their own theories. It goes with the territory.

In my view, Margulis was also wrong about a basic assumption underlying the serial endosymbiosis theory: the idea that the microbial world is a tapestry of many distinct endosymbiotic collaborations. Margulis presented evidence for what she took to be surviving evolutionary intermediates that had never acquired spirochaetes (in the 1967 paper), or mitochondria (in later papers). She linked these purported intermediates specifically with the accumulation of oxygen from photosynthesis. In 1967,

her view was consonant with the latest thinking in geology, cell biology and phylogeny. But decades of work in earth sciences and comparative genomics now shows this view to be far from the truth: the eukaryotes evolved under dysoxic conditions (Knoll et al., 2016), and despite appearances, there are no surviving evolutionary intermediates from early eukaryotic evolution, even though there are very many endosymbiotic interactions among mature eukaryotes (van der Giezen, 2013; Archibald, 2015). That was a shocking turnaround, and could not have been predicted by Margulis or anyone else; even now, many scientists are unwilling to accept the implication that the origin of the eukaryotic cell was a singular event, which occurred just once in four billion years of evolution. By 'origin', I mean the singular endosymbiosis between prokaryotes that set in train the gradualistic evolution of most if not all basal eukaryotic traits.

While we cannot rule out other origins of complex cells (such as the mysterious *Parakaryon myojinensis*; Yamaguchi et al., 2012) it is a conceit to believe that more sophisticated eukaryotes would inevitably outcompete any simpler cells that were evolving greater complexity. That is inconsistent with the abundance of apparent evolutionary intermediates that Margulis and others had pointed to. While they turned out not to be true evolutionary intermediates, they are still ecological intermediates; and they were not outcompeted to extinction by more sophisticated cells. It is equally inconsistent with the fact that neither bacteria nor archaea show any tendency to evolve complex morphological traits found in eukaryotes, and despite their extraordinary metabolic and genetic variation have barely changed in their morphological complexity over four billion years of evolution (Lane 2015). These facts are more consistent with a restrictive bottleneck, in which a fundamental change in cellular structure permitted evolution to take off in a different direction. That structural change probably related to the acquisition of mitochondria, as Margulis argued in her 1967 paper, but had nothing to do with aerobic respiration. Rather, the acquisition of mitochondria restructured genomes in relation to bioenergetic membranes, giving eukaryotes a genomic asymmetry, in which tiny but specialised mitochondrial genomes supported an unprecedented expansion in energy availability per gene.

Acknowledgements

Figure 1
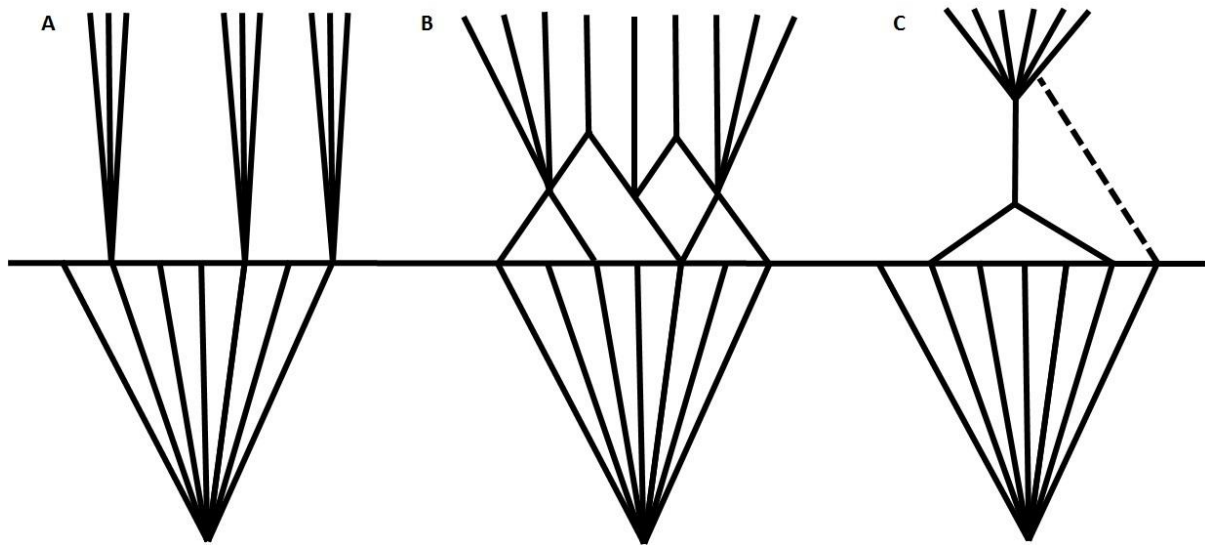
(A) Polyphyletic radiations of complex cells following an environmental bottleneck (horizontal line). Three groups radiate above the line, corresponding to those prokaryotic groups best preadapted to disparate lifestyles, such as phototrophy, osmotrophy, or predation. The complexity of the groups below the line is limited by environmental conditions, such as low marine oxygen concentrations. (B) Polyphyletic radiations of complex cells following an environmental bottleneck that facilitates multiple endosymbioses between diverse prokaryotes, as postulated by the serial endosymbiosis theory. Different groups of complex cells are predicted to arise from distinct endosymbioses in different environments. (C) Monophyletic origin of complex cells, corresponding to a singular endosymbiosis between two prokaryotes. The prokaryotes below the line are constrained by cell structure, not environmental conditions – a restrictive bottleneck – hence successful endosymbioses between prokaryotes are rare events. Here, a rare endosymbiosis gives rise to a long non-branching trunk, in which all basal eukaryotic traits evolved, but no evolutionary intermediates survived, as is observed to be the case. The dotted line represents the later acquisition of plastids in one eukaryotic lineage (algae) but this does not alter the singularity of eukaryotic origins. That applies equally to all other acquisitions of bacterial endosymbionts by 'fully-fledged' eukaryotes.

References

Allen, J.F., 1993. Control of gene expression by redox potential and the requirement for chloroplast and mitochondrial genomes. J. Theor. Biol. 165, 609–631.

Allen, J.F., 2003. The function of genomes in bioenergetic organelles. Phil. Trans. R. Soc. B 358, 19–37.

Allen, J.F., 2017. The CoRR hypothesis for genes in organelles. J. Theoret. Biol. http://doi.org/10.1016/j.jtbi.2017.04.008.

Amend, J.P., Shock, E.L., 2001. Energetics of overall metabolic reactions of thermophilic and hyperthermophilic Archaea and Bacteria. FEMS Microbiol. Rev. 25, 175–243.

Angert, E.R., 2012. DNA replication and genomic architecture of very large bacteria. Annu. Rev. Microbiol. 66, 197–212.

Archibald, J.M., 2009. The puzzle of plastid evolution. Curr. Biol. 19, R81–R88.

Archibald, J.M., 2015. Endosymbiosis and eukaryotic cell evolution. Current Biol. 25, R911–R921.

Barry, E.R., Bell, S.D., 2006. DNA replication in the archaea. Microbiol. Mol. Biol. Rev. 70, 876–887.

Bidle, K.D., Falkowski, P.G., 2004. Cell death in planktonic, photosynthetic microorganisms. Nat. Rev. Microbiol. 2, 643–655.

Blackstone, N.W., 2013. Genetic and energetic aspects of conflict and conflicty mediation. Phil. Trans. R. Soc. B 368, 20120266.

Booth, A., Doolittle, W.F., 2015. Eukaryogenesis, how special really? Proc. Natl. Acad. Sci. USA., 112, 10278–10285.

Boyle, R.A., Clark, J.R., Poulton, S.W., Shields-Zhou, G., Canfield, D.E., Lenton, T.M., 2013. Nitrogen cycle feedbacks as a control on euxinia in the mid-Proterozoic ocean. Nat. Comm. 4, 1533.

Brocchieri, L., Karlin, S., 2005. Protein length in eukaryotic and prokaryotic proteomes. Nucleic Acids Res. 33, 3390–3400.

Canfield, D.E., 1998. A new model for Proterozoic ocean chemistry. Nature 396, 450–453.

Carvalho-Santos, Z., Azimzdeh, J., Pereira-Leal, J.B., Bettencourt-Dias, M., 2011. Tracing the origins of centrioles, cilia, and flagella. J. Cell Biol. 194, 165.

Cavalier-Smith, T., 1987. Eukaryotes with no mitochondria. Nature 326, 332–333.

Cavalier-Smith, T., 1989 Archaebacteria and archezoa. Nature 339, 100–101.

Cavalier-Smith, T., 2002. The phagotrophic origin of eukaryotes and phylogenetic classification of Protozoa. Int. J. Syst. Evol. Microbiol. 52, 297–354.

Cavalier-Smith, T., 2014. The neomuran revolution and phagotrophic origin of eukaryotes and cilia in light of intracellular coevolution and a revised tree of life. Cold Spring Harb. Perspect. Biol. 6, a016006.

Cloud, P.E. Jnr., 1965. Significance of the Gunflint (Precambrian) microflora. Science 148, 27–35.

Crick, F.H.C., 1958. On protein synthesis. Symp. Soc. Exp. Biol. 12, 138–163.

Crick, F.H.C., 1968. The origin of the genetic code. J. Mol. Biol. 38, 367–379.

Daniels, E.W., Breyer, E.P., 1968. Starvation effects on the ultrastructure of amoeba mitochondria. Zeitschrift Zellforschung 91, 159–169.

Dinman, J.D., 2009. The eukaryotic ribosome: current status and challenges. J. Biol. Chem. 284, 11761–11765.

Dobbell, C., 1914. Cytological studies on three species of Amoeba – *A. lacertae* Hartmann, *A. glebae* n.sp., *A. fluvialis* n.sp. Arch. Protistenk. 34, 139.

Elliott, T.A., Gregory, T.R., 2015. What's in a genome? The C-value enigma and the evolution of eukaryotic genome content. Phil. Trans. R. Soc. B 370: 20140331.

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

Errington, J., 2013. L-form bacteria, cell walls and the origins of life. Open Biol. 3, 120143.

Fritz-Laylin, L.K., Prochnik, S.E., Ginger, M.L., Dacks, J.B., Carpenter, M.L., Field, M.C., et al., 2010. The genome of *Naegleria gruberi* illuminates early eukaryotic versatility. Cell 140, 631–642.

Garg, S.G., Martin, W.F., 2016. Mitochondria, the cell cycle, and the origin of sex, via a syncytial eukaryotic common ancestor. Genome Biol. Evol. 8, 1950–1970.

Gerschman, R., Gilbert, D.L., Nye, S.W., Dwyer, P., Fenn, W.O., 1954. Oxygen poisoning and x-irradiation: a mechanism in common. Science 119, 623–626.

Gibor, A., Granick, S., 1964. Plastids and mitochondria: inheritable systems. Science 145, 890–897.

Griese, M., Lange, C., Soppa, J., 2011. Ploidy in cyanobacteria. FEMS. Microbiol. Lett. 323, 124–31.

Hadjivasiliou, Z., Pomiankowski, A., Seymour, R., Lane, N., 2012. Selection for mitonuclear co-adaptation could favour the evolution of two sexes. Proc. R. Soc. B 279, 1865–1872.

Hadjivasiliou, Z., Lane, N., Seymour, R., Pomiankowski, A., 2013. Dynamics of mitochondrial inheritance in the evolution of binary mating types and two sexes. Proc. R. Soc. B 280, 20131920.

Karnkowska, A., Vacek, V., Zubacova, Z., Treitli, S.C., Petrzelvcova, R., Eme, L., et al., 2016. A eukaryote without a mitochondrial genome. Curr. Biol. 26, 1274–1284.

Keeling, P.J., 1998. A kingdom's progress: Archezoa and the origin of eukaryotes. BioEssays 20, 87–95.

Knoll, A.H., Javaux, E.J., Hewitt, D., Cohen, P., 2006. Eukaryotic organisms in Proterozoic oceans. Phil. Trans. R. Soc. B. 361, 1023–1038.

Knoll, A.H., Bergmann, K.D., Strauss, J.V., 2016. Life: the first two billion years. Phil. Trans. R. Soc. B 371, 20150493.

Koonin, E.V., Fedorova, N.D., Jackson, J.D., Jacobs, A.R., Krylov, D.M., Makarova, K.S., et al., 2004. A comprehensive evolutionary classification of proteins encoded in complete eukaryotic genomes. Genome Biol. 5, R7.

Koonin, E.V., 2010. The origin and early evolution of eukaryotes in the light phylogenomics. Genome Biol. 11, 209.

Koonin, E.V., 2012. The Logic of Chance: The Nature and Origin of Biological Evolution. New Jersey, Pearson.

Koonin, E.V., 2015. Origin of eukaryotes from within archaea, archaeal eukaryome and bursts of gene gain: eukaryogenesis just made easier? Phil. Trans. R. Soc. B 370, 20140333.

Lane, N., 2002. Oxygen: The Molecule that Made the World. Oxford, OUP.

Lane, N., 2005. Power, Sex, Suicide: Mitochondria and the Meaning of Life. Oxford, OUP.

Lane N. 2011. Energetics and genetics across the prokaryote-eukaryote divide. Biol. Direct 6, 35.

Lane, N., 2014. Bioenergetic constraints on the evolution of complex life. Cold Spring Harb. Perspect. Biol. 6, a015982.

Lane, N., 2015: The Vital Question: Why is Life the Way it is? London, Profile Books.

Lane, N., Martin, W., 2010. The energetics of genome complexity. Nature 467, 929–934.

Lane, N., Martin, W., 2015. Eukaryotes really are special, and mitochondria are why. Proc. Natl. Acad. Sci. USA. 112, E4823.

Lane, N., Martin, W., 2016. Mitochondria, complexity, and evolutionary deficit spending. Proc. Natl. Acad. Sci. USA. 113, E666.

Lindmark, D.G., Muller, M., 1973. Hydrogenosome, a cytoplasmic organelle of the anaerobic flagellate, *Tritrichomonas foetus*, and its role in pyruvate metabolism. J. Biol. Chem. 248, 7724–28.

Lovelock, J.E., Margulis, L., 1974. Atmospheric homeostasis by and for the biosphere: the Gaia hypothesis. Tellus 26, 2–10.

Lynch, M., Marinov, G.K., 2015. The bioenergetic costs of a gene. Proc. Natl. Acad. Sci. USA. 112, 15690–15695.

Lynch, M., Marinov, G.K., 2017. Membranes, energetics, and evolution across the prokaryote-eukaryote divide. eLife 6, e20437.

Makarova, K.S., Koonin, E.V., 2007. Evolutionary genomics of lactic acid bacteria. J. Bacteriol. 189, 1199–1208.

Mans, B.J., Anantharaman, V., Aravind, L., Koonin, E.V., 2004. Comparative genomics, evolution and origins of the nuclear envelope and nuclear pore complex. Cell Cycle 3, 1612–37.

Margulis, L., Dolan, M.F., Whiteside, J.H., 2005. "Imperfections and oddities" in the origin of the nucleus. Paleobiology 31, 175–191.

Margulis, L. 2006. The phylogenetic tree topples. Am. Sci. 94, 194.

Margulis, L., Sagan, D. 1997. Microcosmos: Folur Billion Years of Microbial Evolution. Oakland, University of California Press.

Margulis, L., Sagan, D., 2002. Acquiring Genomes: A Theory of the Origins of Species. New York, Basic Books.

Martijn, J., Ettema, T.J.G., 2013. From archaeon to eukaryote: the evolutionary dark ages of the eukaryotic cell. Biochem. Soc. Trans. 41, 451–457.

Martin, W., 1999. Mosaic bacterial chromosomes: a challenge en route to a tree of genomes. BioEssays 21, 99–104.

Martin, W., Koonin, E.V., 2006. Introns and the origin of nucleus–cytosol compartmentalization. Nature 440, 41–45.

Martin, W., Müller, M., 1998. The hydrogen hypothesis for the first eukaryote. Nature 392, 37–41.

Martin, W.F., Roettger, M., Ku, C., Garg, S.G., Nelson-Sathi, S., Landan, G. Late mitochondrial origin is pure artefact. bioRxiv doi: https://doi.org/10.1101/055368.

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. Biol. Chem. 382, 1521–1539.

Maynard Smith, J., Szathmary, E., 1995. The Major Transitions of Evolution. Oxford, OUP.

McClutcheon, J.P., Moran, N.A., 2012. Extreme genome reduction in symbiotic bacteria. Nat. Rev. Microbiol. 10, 13–26.

McCutcheon, J.E., von Dohlen, C.D., 2011. An independent metabolic patchwork in the nested symbiosis of mealybugs. Curr. Biol. 21, 1366–1372.

McInerney, J.O., Martin, W., Koonin, E.V., Allen, J.F., Galperin, M.Y., Lane, N., Archibald, J.M., Embley, T.M., 2011. Planctomycetes and eukaryotes: a case of analogy not homology. BioEssays 33, 810-817.

Mendell, J.E., Clements, K.D., Choat, J.H., Angert, E.R., 2008. Extreme polyploidy in a large bacterium. Proc. Natl. Acad. Sci. USA. 105, 6730–6734.

Mentel, M., Martin, W., 2008. Energy metabolism among eukaryotic anaerobes in light of Proterozoic ocean chemistry. Phil. Trans. R. Soc. B. 363, 2717–29.

Miller, S.L., 1953. A production of amino acids under possible primitive Earth conditions. Science 117, 528–29.

Minerdi, D., Bianciotto, V., Bonfante, P., 2002. Endosymbiotic bacteria in mycorrhizal fungi: from their morphology to genomic sequences. Plant Soil 244, 211–219.

Müller, M., Mentel, M., van Hellemond, J.J., Henze, K., Woehle, C., Gould, S.B., Yu, R.-Y., van der Giezen, M., Tielens, A.G.M., Martin, W.F., 2012. Biochemistry and evolution of anaerobic energy metabolism in eukaryotes. Microbiol. Mol. Biol. Rev. 76, 444–495.

Murphy, M.P., 2009. How mitochondria produce reactive oxygen species. Biochem. J. 417, 1–13.

Nowack, E.C., Grossman, A.R., 2012. Trafficking of proteins into the recently established photosynthetic organelles of *Paulinella chromatophora*. Proc. Natl. Acad. Sci. USA. 109, 5340–5345.

Nývltová, E., Stairs, C.W., Hrdý, I., Rídl, J., Mach, J., Pačes, J., Roger, A.J., Tachezy, J., 2015. Lateral gene transfer and gene duplication played a key role in the evolution of *Mastigamoeba balamuthi* hydrogenosomes. Mol. Biol. Evol. 32, 1039–1055.

Patel, B.H., Percivalle, C., Ritson, D.J., Duffy, C.D. and Sutherland, J.D., 2015. Common origins of RNA, protein and lipid precursors in a cyanosulfidic protometabolism. Nat. Chem. 7, 301–07.

Pittis, A.A., Gabaldon, T., 2016. Late acquisition of mitochondria by a host with chimaeric prokaryotic ancestry. Nature 531, 101-104.

Qiu, H., Price, D.C., Weber, A.P.M., Facchinelli, F., Yoon, H.S., Bhattacharya, D., 2013. Assessing the bacterial contribution to the plastid proteome. Trends Plant Sci. 18, 680–687.

Radzvilavicius, A.L., Blackstone, N.W., 2015. Conflict and cooperation in eukaryogenesis: implications for the timing of endosymbiosis and the evolution of sex. J. R. Soc. Interface 12, 20150584.

Sagan, L., 1967. On the origin of mitosing cells. J. Theoret. Biol. 14, 225–274.

Schulz, H.N., 2006. The genus Thiomargarita. Prokaryotes 6, 1156–1163.

Schulz, H.N., Jorgensen, B.B., 2001. Big bacteria. Annu. Rev. Microbiol. 55, 105–137.

Sousa, F.L., Neukirchen, S., Allen, J.F., Lane, N., Martin, W.F., 2016. Lokiarchaeon is hydrogen dependent. Nature Microbiol 1, 16034.

Spang, A., Saw, J.H., Jørgensen, S.L., Zaremba-Niedzwiedzka, K., Martijn, J., Lind, A.E., van Eijk, R., Schleper, C., Guy, L., Ettema, T.J., 2015. Complex archaea that bridge the gap between prokaryotes and eukaryotes. Nature 521, 173–79.

Staub, E., Fiziev, P., Rosenthal, A., Hinzmann, B., 2004. Insights into the evolution of the nucleolus by an analysis of its protein domain repertoire. Bioessays26: 567–581.

Vafai, S.B., Mootha, V.K., 2012. Mitochondrial disorders as windows into an ancient organelle. Nature 491, 374–383.

van der Giezen, M., 2013. Evolution: One thread to unite them all. Curr. Biol. 23, R679–R681.

Vellai, T., Takacs, K., Vida, G., 1998. A new aspect to the origin and evolution of eukaryotes. J. Mol. Evol. 46, 499–507.

von Dohlen, C.D., Kohler, S., Alsop, S.T., McManus, W.R.. 2001. Mealybug beta-proteobacterial endosymbionts contain gamma-proteobacterial symbionts. Nature 412, 433–436.

Ward, R.J., Clements, K.D., Choat, J.H., Angert, E.R., 2009. Cytology of terminally differentiated Epulopiscium mother cells. DNA Cell. Biol. 28, 57–64.

Wallin, I.E., 1927. Symbionticism and the Origin of Species. Baltimore, Williams and Wilkins.

Wernegreen, J.J., 2012. Endosymbiosis. Current Biol. 22, R555–R561.

Wilson, E.B., 1925. The Cell in Development and Heredity. New York, Macmillan.

Williams, T.A., Foster, P.G., Cox, C.J., Embley, T.M., 2013. An archaeal origin of eukaryotes supports only two primary domains of life. Nature 504, 231–236.

Woese, C.R., 1967. The Genetic Code: The Molecular Basis for Gene Expression. New York, Harper and Row.

Woese, C.R, Fox, G.E., 1977. Evolution Phylogenetic structure of the prokaryotic domain: The primary kingdoms. Proc. Natl. Acad. Sci. USA. 74, 5088–5090.

Wujek, D.E., 1979. Intracellular bacteria in the blue-green-alga *Pleurocapsa minor.* Trans. Am. Micros. Soc. 98, 143–145.

Yang, S., Doolittle, R.F., Bourne, P.E., 2005. Phylogeny determined by protein domain content. Proc. Natl Acad. Sci. USA. 102, 373–378.

Yamaguchi, M., Mori, Y., Kozuka, Y., Okada, H., Uematsu, K., Tame, A., Furukawa, H., Maruyama, T., Worman, C.O., Yokoyama, K., 2012. Prokaryote or eukaryote? A unique microorganism from the deep sea. J. Electron Microsc. (Tokyo). 61, 423–431.

Yutin, N., Wolf, M.Y., Wolf, Y.I., Koonin, E.V., 2009. The origins of phagocytosis and eukaryogenesis. Biol. Direct 4, 9.

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 doi: 10.1038/nature21031

Zouni, A., Witt, H.T., Kern, J., Fromme, P., Krauss, N., Saenger, W.,Orth, P., 2001. Crystal structure of photosystem II from *Synechococcus elongatus* at 3.8 Å resolution. Nature 409, 739–743.

Zubacova, Z., Cimburek, Z. & Tachezy, J., 2008. Comparative analysis of trichomonad genome sizes and karyotypes. Mol. Biochem. Parasitol. 161, 49–54.

Zuckerkandl, E., Pauling, L., 1965. Molecules as documents of evolutionary history. J. Theoret. Biol. 8, 357–366.