PERSPECTIVES

OPINION

Mitochondrial and plastid evolution in eukaryotes: an outsiders' perspective

Jeferson Gross and Debashish Bhattacharya

Darwin200

Abstract | The eukaryotic organelles mitochondrion and plastid originated from eubacterial endosymbionts. Here we propose that, in both cases, prokaryote-to-organelle conversion was driven by the internalization of host-encoded factors progressing from the outer membrane of the endosymbionts towards the intermembrane space, inner membrane and finally the organelle interior. This was made possible by an outside-to-inside establishment in the endosymbionts of host-controlled protein-sorting components, which enabled the gradual integration of organelle functions into the nuclear genome. Such a convergent trajectory for mitochondrion and plastid establishment suggests a novel paradigm for organelle evolution that affects theories of eukaryogenesis.

A hallmark of eukaryotic cells is the presence of a nucleus and the bioenergetic organelles — the mitochondrion and the plastid, which originated from Gram-negative bacterial progenitors that once resided inside a eukaryotic 'host' cell as endosymbionts and were converted over time into bona fide organelles[1–6]. It is believed that the mitochondrion originated through endosymbiosis from an alphaproteobacterium approximately two billion years ago[1–6], but the nature of the host remains uncertain and lies at the heart of an ongoing debate about eukaryogenesis[3]. One line of reasoning argues that an endomembrane system was a prerequisite for phagocytic capture of the alphaproteobacterium, implying that the host of the mitochondrial forerunner was a primitive eukaryote[3,6]. The opposing view is that the eukaryotic stem lineage was founded by the association of the alphaproteobacterial endosymbiont with a methanogenic archaeon, with the endosymbiont exerting a prominent influence on the emergence of the endomembrane system and the nuclear envelope in the methanogenic host[7,8].

Despite these divergent opinions about mitochondrial origin, there is broad agreement that the plastid originated approximately 500 million years later in a bona fide mitochondriate eukaryotic host through a cyanobacterial primary endosymbiosis[1,2,9]. Although the phylogenetic provenance of the cells (if not the nature of the host) that gave rise to the mitochondrion and the plastid has been clarified, a satisfactory explanation of how the prokaryote-to-organelle conversion (organellogenesis) occurred is lacking. However, it is clear that this process led to a strict dependence of the organelles on the nucleus[1,2,4–6,10–12]. This is substantiated by the fact that over 2,000 proteins that function in modern day mitochondria and plastids are encoded in the nuclear genome, synthesized by cytosolic ribosomes and then transported into the endosymbiont-derived compartments[13,14]. Such a nucleus-to-organelle flow of information constitutes the essential property of organellogenesis that resulted from the eukaryotic host progressively assuming control over the biogenesis of the captured prokaryotic cells.

Nuclear dominance over bioenergetic organelles is probably the result of continuous selective pressure on the host to optimize energy production from alphaproteobacterial oxidative phosphorylation and cyanobacterial oxygenic photosynthesis[1,2,11,12]. Therefore, innovations that increased the efficiency of energy production and its extraction from the endosymbionts were fixed over time in the nuclear genome. In this context, it is likely that the genomes of the mitochondrial and plastid forerunners were degenerating owing to Muller's ratchet[10,11,15], resulting in a decrease in bioenergetic performance. To maintain and further improve the physiological competence of the endosymbionts, the mitochondrial and plastid hosts were under selective pressure to take over the functions of the captive prokaryotes. This led to the establishment of nuclear genes that encode products that optimize the fitness of endosymbiont-derived compartments[11,12,16–18]. These genes were recruited by endosymbiotic gene transfer (EGT)[16], by the *de novo* generation of genes[11,12], by horizontal gene transfers (HGT)[18] and by co-option of pre-existing host functions[17]. These processes resulted in massive endosymbiont (that is, organellar) genome reduction (to 20–200 genes, data from NCBI organelle genome resources) owing to losses of dispensable functions or replacement of functions by nuclear loci[10,11].

Nuclear control of endosymbiont activities necessitated the evolution of protein-sorting systems to recognize and relocate nuclear-encoded polypeptides into the nascent organelles[1,2,4–6,11,12]. Such primitive protein-sorting systems correspond to the ancestors of the Toc and Tic protein translocons in modern-day plastids[19–21], and the sorting and assembly machinery (SAM), the Tom and Tim23 translocons and the solute carrier Tim22 insertase in modern-day mitochondria[4,22,23] (the putative phylogenetic origins of the subunits of complexes is presented in Supplementary information S1 (table)). Traditional views usually assume that these components arose to target nuclear-encoded proteins that contain presequences directly into the interior of the endosymbionts (that is, the prokaryotic

cytosol) to support increasing EGT in the early stages of organellogenesis[4,5,12,19,20,24,25]. This 'insiders' paradigm' is also a characteristic of alternative models that assume the initial target for organelle establishment was the random insertion of eukaryote-derived solute transporters in the inner membrane (IM) of the endosymbionts[6]. As Gram-negative bacteria are shielded by four external layers (lipopolysaccharides, the outer membrane (OM), peptidoglycan and the IM)[26], these models fail to explain how nuclear-encoded molecular components gained access to the interior of prokaryotic cells. By taking an 'outsiders' perspective', we propose instead that the foundation for mitochondrion and plastid evolution was laid in the OM of the captured endosymbionts because of its greater accessibility to host factors in the cytoplasm. Only later could host-directed components systematically enter the intermembrane space (IMS), then gain access to the IM, and finally reach the organelle interior. Such a putative outside-to-inside trajectory provides a straightforward and unanticipated perspective on organelle evolution by implying that both the mitochondrion and the plastid were established in stages in a convergent manner.

The story begins in the OM

A minimal nuclear-encoded mitochondrial system. We propose that mitochondrial and plastid organellogenesis was initiated by the respective hosts assuming control over the assembly of β-barrel proteins in the OM of the endosymbionts (BOX 1). This occurred when *Omp85* genes were established by EGT in the nuclear genome of the eukaryotic hosts, giving rise to the ancestors of *Sam50* and *Toc75* (REFS 4,27) (BOX 1). In the case of the mitochondrion we propose that, parallel to the establishment of *Sam50* in the nucleus, another component that was probably derived from a bacterial porin gene evolved in the host nuclear genome. Its product was expressed, synthesized by cytosolic ribosomes and targeted to the OM of the endosymbiont. This pore-forming component progressively acquired the property to function as a receptor and a translocation channel to dock and import itself and Sam50 across the external alphaproteobacterial membrane (FIG. 1a). The incipient OM pore represents the ancestor of the β-barrel protein Tom40, which is the core of the Tom translocon[4]. Each newly translocated Tom40 and Sam50 molecule that gained access to the endosymbiont periplasm could potentially be recognized by endogenous prokaryotic chaperones (for example, Skp and SurA in *Escherichia coli*) and delivered to the pre-existing alphaproteobacterial Omp85 and inserted in the OM of the endosymbiont[26]. Once pools of Sam50 accumulated in the OM of the endosymbiont, this component could progressively assume control over its own assembly and that of Tom40 (FIG. 1a), as observed in modern-day mitochondria[22,23]. We postulate that this constituted a minimal organized system of nuclear-encoded factors that were able to target and assemble themselves in the OM of the endosymbiont. The striking aspect of this initial sorting system was its property of self-sustainability and amplification (BOX 2), which was probably a feature required to drive its own establishment.

A minimal nuclear-encoded plastid system. In plastids, the product of the cyanobacterial *Omp85* gene that was established in the nucleus progressively acquired the ability to dock and translocate itself across the OM, resulting in the ancestor of Toc75 (FIG. 1b). Initially, incoming Toc75 in the endosymbiont periplasm was taken by pre-existing chaperones and assembled by cyanobacterial Omp85. As observed for ancestral mitochondrial Sam50, primitive Toc75 maintained its prokaryotic Omp85 property and, once inserted in the OM, could drive the assembly of incoming pools of Toc75. This was the origin of a self-sustaining loop of translocation and assembly around Toc75 (BOX 2). It is interesting that a similar biogenetic path is retained by modern-day Toc75 in plant chloroplasts in which plastid-targeted Toc75 precursors are translocated into the IMS of the organelle by pre-existing pools of Toc75 and are assembled in the OM[28].

The validity of the ideas discussed above will depend on experimental data that link the function of Toc75-related proteins in plants and algae to the insertion of β-barrel proteins in the OM of plastids. It has been

Box 1 | **Priming the prokaryotic outer membrane for organellogenesis**

Whereas proteins localized to the inner membrane (IM) of Gram-negative bacteria are folded according to an α-helical structural motif, the vast majority of proteins contained in the outer membrane (OM) are of the β-barrel type[26]. The OMs of plastids and mitochondria are descended from the OMs of their prokaryotic Gram-negative progenitors, and retain the ability to assemble β-barrel proteins[4,29,31,32]. The two best examples of this type of OM component in mitochondria are Sam50 and the protein-conducting channel of the Tom translocon, Tom40 (REFS 4,22,74). Similarly, the pore of the Toc translocon in plastids, Toc75, is a β-barrel protein[20,27,29]. In Gram-negative bacteria, β-barrel components are assembled by the outer membrane protein of 85 kDa (Omp85), which is a β-barrel protein pore capable of self-assembly[26]. In mitochondria, the integration of this class of protein in the OM is carried out by the sorting and assembly machinery (SAM) complex — the central component of this complex is Sam50, a homologue of alphaproteobacterial Omp85 (REFS 4,23,26). Curiously, Toc75 is also a homologue of cyanobacterial Omp85 (REFS 27,29). The fact that both mitochondria and plastids conserve essential features of Gram-negative OMs (that is, the presence of β-barrel proteins and functional Omp85 homologues) is a remarkable example of convergence in the evolution of plastids and mitochondria. We suggest this did not occur by chance but rather reflects the fact that the establishment of the first molecular system of host-encoded factors in the OM of the endosymbionts was in the context of β-barrel protein assembly.

But how did host-derived components gain access to the OM? We assume that the Gram-negative bacterial forerunners of the mitochondrion and the plastid were free inside the host cell (for example, by escaping digestion) with their OMs directly exposed to the cytoplasm. It is conceivable that permeability of the OM of the captured prokaryotes occurred as a consequence of a long-term endosymbiotic relationship that allowed access to host-derived cytosolic proteins. This may have been caused by, for example, the loss of structure of the OM lipopolysaccharides that is observed among modern-day prokaryotic endosymbionts[75]. Alternatively, as in mitochondria, plastids and some endosymbiotic bacteria in eukaryotes[75–78], it is conceivable that pools of glycerolipids derived from the ancestral host endoplasmic reticulum (for example, phosphatidylcholine) might have systematically gained access to the OM of the plastid and mitochondrial precursors and thereby increased membrane permeability. Another intriguing possibility is that access of host-encoded factors to the OM may have been facilitated by delivery of proteins through vesicular systems derived from the secretory pathway[1] or, alternatively, by co-translational targeting of nascent polypeptides coordinated by cytosolic ribosomes attached to the endosymbiont outermost lipid layer[12]. In any case, we suggest that the combination of an acquired accessibility to the OM of the endosymbionts and the putative operation of a primitive protein delivery system helped set the stage for the evolution of an initial organized system of nuclear-encoded molecular factors that were able to act in the OM of the plastid and mitochondria forerunners.

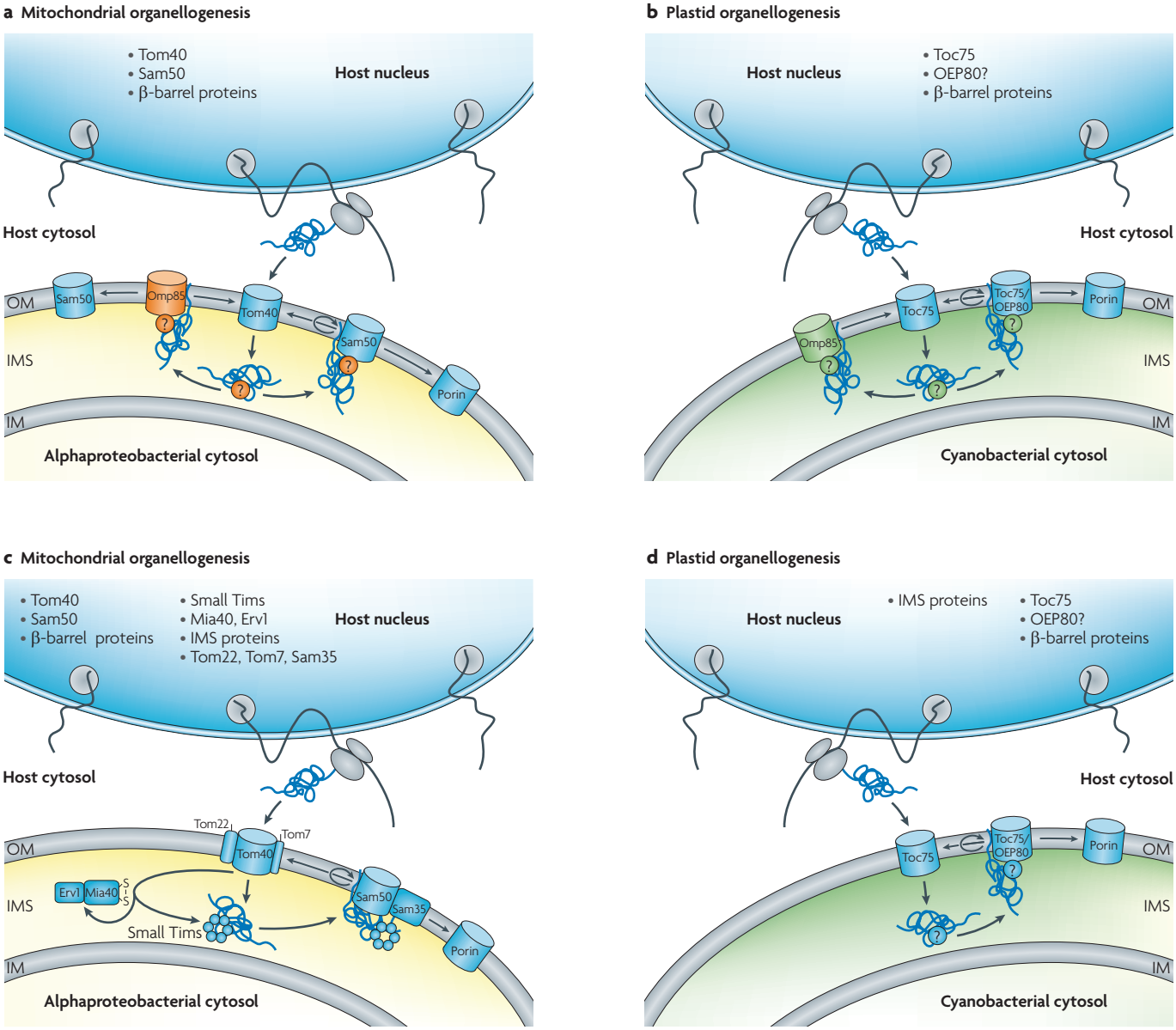


Figure 1 | **Putative early stages in mitochondrial and plastid organellogenesis.** Molecular factors are either encoded in the nucleus (blue) or in the alphaproteobacterial (yellow) or cyanobacterial (green) chromosome. Question marks indicate currently unknown components. **a**,**b** | Genes encoding Tom40, Sam50 (**a**), Toc75 and possibly OEP80 (**b**) were relocated to the nuclear genome of the respective eukaryotic hosts, and the encoded mRNA translated by cytosolic ribosomes. The resulting proteins were directed across the Tom40 pore in the alphaproteobacterial outer membrane (OM) (**a**) or the Toc75 channel in the cyanobacterial OM (**b**), then protected by bacterial periplasmic chaperones in the intermembrane space (IMS), and assembled in the OM by endogenous Omp85. Accumulated OM pools of Sam50 (**a**) and Toc75 (or OEP80) (**b**) increasingly assumed control over the assembly of nuclear-encoded proteins, enabling the establishment in the nucleus of new genes for organelle-destined β-barrel proteins (for example, porins). **c**,**d** | Novel nuclear-encoded factors gained access to the IMS, replaced the endogenous components, and supported the biogenesis of β-barrels by Sam50 in the alphaproteobacterial OM (**c**) and Toc75 (or OEP80) in the cyanobacterial OM (**d**). Erv1 and Mia40 (**c**) were introduced during mitochondrial evolution as a disulphide relay system for newly established IMS proteins, including the ancestor of the small Tim chaperones and Mia40–Erv1. At this stage, the subunits Tom22 and Sam35 (**c**) could have been added as receptors for β-barrel proteins[81,74] and Tom7 as a negative regulator of Tom40 biogenesis[23]. IM, inner membrane.

proposed that the biogenesis of β-barrel proteins in the chloroplast OM relies on the outer envelope protein of 80 kDa (OEP80), which is a Toc75 paralogue[29,30]. If this hypothesis is correct, then two scenarios may explain the diversity of Toc75-related proteins in plants and algae. In the first scenario, a duplication of the primitive *Toc75* gene occurred during evolution. The paralogue OEP80 became predominantly involved in the assembly of β-barrel proteins in the plastid OM, whereas Toc75 specialized in translocation of proteins across the OM of the organelle (FIG. 1b). In the second scenario, two paralogues of cyanobacterial *Omp85* were already established in the nucleus of the host in the early stages of organellogenesis. These two loci evolved to encode a self-sustaining loop around the protein-conducting pore Toc75 and the β-barrel assembly factor OEP80 (FIG. 1b).

A nuclear-encoded 'Trojan horse'. Tom40–Sam50 and Toc75–Toc75 were established as minimal systems in the OM of the endosymbionts. From here, they could progressively displace endogenous Omp85 in complexes with their native auxiliary factors (for example, the YfgL, NlpB and YfiO interactors of the *E. coli* Omp85)[26] and assume control over the assembly of β-barrel proteins encoded by the prokaryotes[31]. This precipitated the loss of the original *Omp85* gene from the chromosome of the endosymbionts. In addition, prokaryote β-barrel proteins (for example, porins) could be established in the nucleus by EGT or HGT and thereafter be imported through the primitive Tom40 and Toc75 pores and assembled by Sam50 and Toc75 in the OM of the incipient mitochondrion and plastid, respectively[31,32] (FIG. 1a,b). These porin-type components would potentially have formed the first layer of metabolic flow regulation between the host cytosol and the endosymbiont interior. These events illustrate the principle that genetic integration of molecular components acting in a given endosymbiont compartment (for example, the OM) follows the establishment of 'pioneering' protein-sorting factors in that compartment.

The IMS as a new host space

The establishment of the pores Tom40 and Toc75 in the OM of mitochondria and plastids, respectively, allowed newly evolved nuclear-encoded proteins to gain access to the endosymbiont periplasm (the organellar IMS). The best example of this phenomenon is provided by the ancestor of the small Tim chaperone complexes, Tim9–Tim10 and Tim8–Tim13, in mitochondria[22,33]. These paralogues originated in the nucleus of the host from a single progenitor[33] that we suggest evolved in the context of the initial self-sustaining loop in the OM to improve the efficiency of Sam50-driven β-barrel protein assembly (FIG. 1c; BOX 2). This function is still maintained by small Tim chaperones in modern-day mitochondria that protect hydrophobic β-barrel proteins (such as Tom40 and Sam50) from self-aggregation once they traverse the Tom40 pore and gain access to the IMS[22,23,33]. The Tim chaperones also escort these sorting substrates to the Sam50 complex (FIG. 1c). In plastids, Toc75 is also inserted from the IMS[28], which suggests that the prokaryotic paradigm for β-barrel assembly from the inside might have been recapitulated in both mitochondrion and plastid evolution. Analogous to the mitochondrion, it is possible that a host-controlled chaperone system may also have evolved in the IMS of plastids to support the assembly of β-barrel proteins by ancestral Toc75 (FIG. 1d).

Interestingly, the small Tim proteins in the mitochondrial IMS and the Skp and SurA components in the periplasm of many Gram-negative bacteria are analogous ATP-independent chaperone systems[34,35]. Also, the structure of the complexes formed by small Tim components resembles the architecture of the Skp complex, although no relationship in amino acid sequence exists between these chaperones[35]. Even the binding specificity of Tim10 and SurA for hydrophobic substrates has been shown to be remarkably similar[34]. In addition, the redox relay system that exists in the IMS of mitochondria to import (by a folding trap) and lead to the maturation of disulphide-containing proteins, such as small Tim proteins, is analogous to the DsbB–DsbA pathway in the periplasm of Gram-negative bacteria, except that the reciprocal components existing in mitochondria, Erv1 and Mia40 (FIG. 1c), are of eukaryotic origin[36]. All of these analogies support the idea that newly evolved nuclear-encoded factors gained access to the endosymbiont periplasm, displaced the endogenous components and took over their functions and respective substrates. This host-controlled IMS system served as a platform for a major downstream event in organelle evolution.

Reprogramming the IM permeome

Eukaryotic solute carriers in organelles. Mitochondria and plastids are metabolically active compartments that routinely exchange

Box 2 | The role of autocatalytic networks in organellogenesis

The concept of autocatalytic networks is often applied in origin-of-life models to describe the replication behaviour of a given set of two or more components that are capable of catalysing their own synthesis[79]. The biogenetic interdependence of Tom40 and Sam50 (REFS 23,75) can be described as an autocatalytic system because the presence of Tom40 and Sam50 enhances the import and assembly of more of these proteins. The advantage of an autocatalytic organization resides in its self-sustainability, as well its evolvability. Each mutation that enhances the function of a component affects the fitness of the whole autocatalytic network[79]. Through this mechanism, a weak selective pressure on the function of Sam50 would favour the fixation of mutations that improve the translocation activity of Tom40, and vice versa.

Interestingly, the amplificatory property of the Tom40–Sam50 autocatalytic loop probably required the early establishment of a negative regulator. This might have been Tom7 (FIG. 1c), a ubiquitously conserved subunit of the Tom complex that in yeast mitochondria exerts negative control over the turnover of Tom40 assembly in the Tom complex[23]. Another inherent aspect of an autocatalytic organization is its expansibility. This occurs when a new component is introduced to the network as a by-product of one of the previous components and reinforces the biogenesis of the upstream factors. One example is the establishment of the ancestral small Tim in the intermembrane space (IMS), which stimulated the assembly of Sam50 and Tom40 in the outer membrane (OM)[23,33]. As a result of this, accumulation of Tom40 facilitated access to the IMS by an increasing pool of small Tim chaperones. In a similar manner, the introduction of the Mia40–Erv1 system provided a relay of disulphide bonds to the cysteines of the small Tim proteins[36,80], thereby maturating and trapping functional Tim chaperones in the IMS. Consequently, more Sam50 and Tom40 were assembled, stimulating the translocation of Mia40 and Erv1 by Tom40. Curiously, in plants and metazoans the entrapment of Mia40 and Erv1 in the IMS depends on disulphide bonds relayed by previously established Mia40–Erv1 (REF. 81), thereby strengthening the whole biogenetic cycle. The reinforcing loop mechanism spread to the inner membrane (IM) as well. It is not by chance that Tim22 insertion is autocatalytic and dependent on the function of small Tim proteins. In turn, Tim22 assembles Tim23 and Tim17 (REFS 22,23). Curiously, the fungal Mia40 is a typical single transmembrane domain substrate of the Tim23 complex[80]. Therefore, the entire mitochondrial protein-sorting system was gradually built up on a multi-component autocatalytic network.

The hypothetical bioenergetic interdependence of the Toc75 translocator and the Toc75 β-barrel assembly factor also included a true autocatalytic loop in the primitive plastid. The observation that Toc75 is the only OM protein that relies on a transit peptide (TP)[28] might provide a further example of this phenomenon. In parallel to the establishment of Toc75-mediated translocation of TP-containing proteins, the ancestral Toc75 acquired a TP, thereby conditioning its biogenesis to its new function. The translocation of TP-containing substrates in the IM relies on Tic110, which self-catalyses its own translocation[68,69] and is possibly involved in its further integration in the IM from the organelle interior[67,69]. In addition, the precursor of Toc75 depends on its positioning at the Tic110 pore for processing of the TP before assembly in the OM[28]. Therefore, the entire apparatus for TP translocation was putatively organized as a self-reinforcing system.

with the cytosol a plethora of small metabolites such as metal ions, amino acids, carbohydrates, nucleotides and cofactors. The regulation of metabolic flow between the organelles and the cytosol relies on an array of substrate-specific transporters located in the IM of mitochondria and plastids — the IM permeome[37]. This is also the case in Gram-negative bacteria, in which permeability between the prokaryotic cytosol and the external environment is regulated by IM transporters[17]. Surprisingly, phylogenetic studies show that solute carriers in the IM of mitochondria and plastids are largely of eukaryotic rather than prokaryotic provenance[11,12,17,38]. Mitochondrial carriers are a specialized family of six α-helix-containing proteins that are integral to the IM and that evolved in the host nucleus[38], whereas in plastids the majority of solute carriers were co-opted from transporters that already existed in the vacuolar membrane system of the host[17]. These data show a major transition in organellogenesis that occurred when the hosts replaced all previous IM solute transporters of the endosymbionts with eukaryote-derived nuclear-encoded solute carriers (FIG. 2a,b). Since then, the eukaryote hosts acquired the ability to regulate the function and physiology of the organelle to best suit their needs. Examples of such newly evolved selective solute carriers are the ADP/ATP carrier in mitochondria that allowed the host to tap into energy produced by oxidative phosphorylation[11,12,38], and the triose-phosphate/phosphate translocator in plastids that made the flow of carbohydrates generated by photosynthesis possible[39].

Requirements for permeome reprogramming. The systematic insertion of eukaryote solute carriers in the IM of the endosymbionts was crucial to the evolution of organelles from captured prokaryotes. This raised the possibility that the onset of mitochondrial and plastid organellogenesis occurred directly in the IM by random insertions of incipient solute transporters[6]. However, as the topogenesis of integral proteins is usually a catalysed process[22,40], it is more likely that primitive insertases evolved early in the nascent organelles to fulfil this task. One example is Tim22 in the mitochondrion, a protein responsible for insertion of the mitochondrial carriers Tim17 and Tim23, as well as itself, into the IM[22,23,41] (FIG. 2a). Tim22 is broadly conserved among eukaryotes and, in contrast to previous suggestions[41], seems to lack a homologue in bacteria (Supplementary information S1 (table)). We propose that Tim22 evolved in the eukaryotic host as a factor to drive the reprogramming of the alphaproteobacterial IM permeome. An important observation is that Tim22 and its substrates are imported across the Tom40 channel[22,23]. In addition, before insertion in the IM, hydrophobic Tim22 and the mitochondrial carriers need to be escorted through the aqueous IMS by a complex of small Tim chaperones (Tim9–Tim10) similar to the complex that was previously found to support β-barrel biogenesis by Sam50 in the OM[22,23] (FIG. 2a). These data indicate that the OM and IMS systems that were in place in earlier stages of organellogenesis played a fundamental part in the evolution of the organized insertion of solute carriers in the IM of the nascent mitochondrion by Tim22.

Harnessing the respirasome

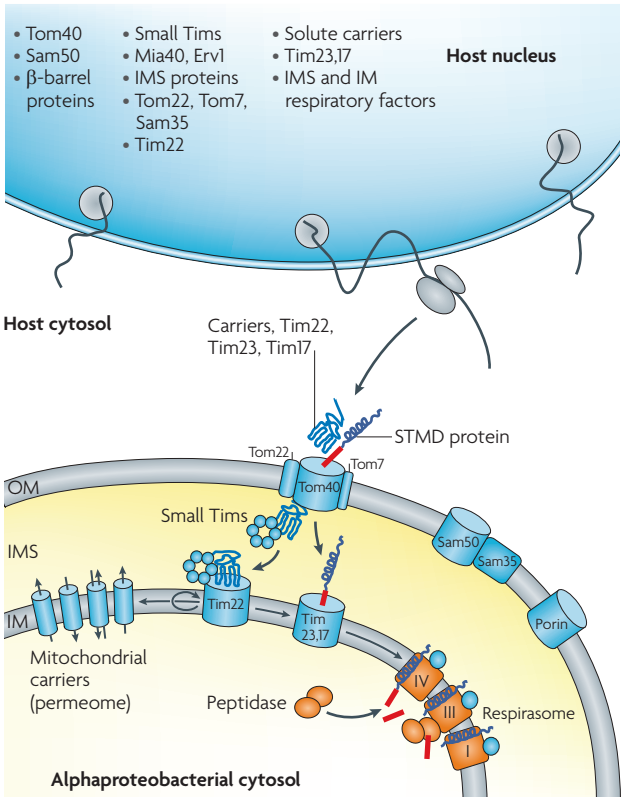
N-terminal presequences before matrix targeting. Tim23 and Tim17, which form the core of the Tim23 translocase[22,23] in the IM of mitochondria, were the next players to be introduced in this evolutionary 'saga' (FIG. 2c). These proteins are phylogenetically related to Tim22 and are integrated in the IM by the Tim22 insertase[4,22,41]. These observations suggest that *Tim23* and *Tim17* might have been generated by a duplication involving an ancestral *Tim22* gene, with family members evolving to perform two different protein pore functions (as the Tim22 carrier insertase and as a member of the Tim23 complex). We propose that the Tim23 complex originally evolved to insert proteins anchored in the IM by an N-terminal, α-helical single transmembrane domain (STMD)[42,43] (FIG. 2a). Proteins inserted by Tim23 typically have a positively charged and cleavable N-terminal extension that is traditionally referred to as a matrix targeting sequence[22,44]. We suggest that such N-terminal presequences evolved to increase the overall positive charge in the N-terminus of the Tim23 substrates that, under the influence of the IM electrochemical potential (negatively charged in the matrix side of the IM), helped the presequences to sink into the Tim23 channel and place the STMD in the plane of the IM[42,45,46]. Therefore, presequences might have initially evolved not as topogenic signals but as devices to enhance the IM integration of STMD proteins (BOX 3) in a stage of organellogenesis when translocation into the matrix was not yet mechanistically feasible.

Is Tim23 the respirasome translocase? The I (NADH:ubiquinone oxidoreductase), III (cytochrome bc_1) and IV (cytochrome *c* oxidase) complexes of the mitochondrial respiratory chain tend to be grouped in multi-complex assemblies often called the respirasome[47] (FIG. 2a). In most mitochondria a core of usually 20 components derived from the alphaproteobacterial ancestor confer the catalytic properties of the respiratory enzymes; that is, the transfer of electrons to oxygen and the pumping of protons into the IMS[48–52]. Surprisingly, in eukaryotes the mitochondrial complexes I, III and IV contain several additional small subunits that are not present in bacteria[48–52]. In the mammalian respirasome there are ~48 of these supernumerary subunits, ~24 of which are STMD-type proteins[48–52]. With some exceptions, the function of these new small components is linked to the stability and, in particular, to the assembly of the complexes[48,51,52]. We propose that once the host gained access to the IM, a novel strategy to exert external control over the bioenergetic performance of the endosymbiont evolved through the establishment of Tim23: the introduction of STMD subunits in the IM sector of the respiratory chain (FIG. 2a).

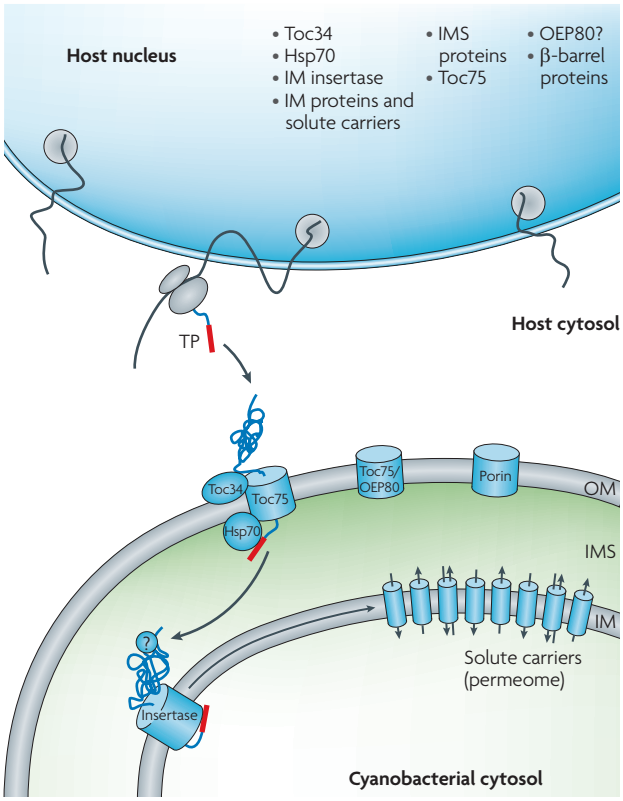
Because all respirasome subunits were encoded in the alphaproteobacterial chromosomes, the new principle was to break the endogenous assembly of the respiratory complexes in sequential steps — the transitions between these steps depended on the coordinated addition and action of the nuclear-encoded small subunits[48,51,52]. The expression of these subunits could have been fine-tuned under nuclear regulatory circuits that evolved to sense oxygen[53,54]. Thereby, newly evolved small supernumerary subunits could dictate the rhythm of assembly and disassembly of the respirasome to optimize bioenergetic performance and minimize production of harmful reactive oxygen species[53]. At least nine STMD subunits in mammals are conserved in plants and fungi[48–52]. In addition, components involved in the biogenesis of the respiratory cofactors (for example, Cox11 and Sco1) and the catalytic subunit of complex III, cytochrome c_1, are STMD proteins[51,53]. All of these topologically simple polypeptides could potentially have been encoded in the nucleus following evolution of the Tim23 insertase. Because sorting to the IMS was already functional, soluble supernumerary subunits could also aggregate to the part of the respirasome exposed to the IMS (FIG. 2a).

Remarkably, the β- and α-subunits of the matrix processing peptidase (MPP) that is involved in cleavage of the Tim23 presequences are homologues of, respectively, the

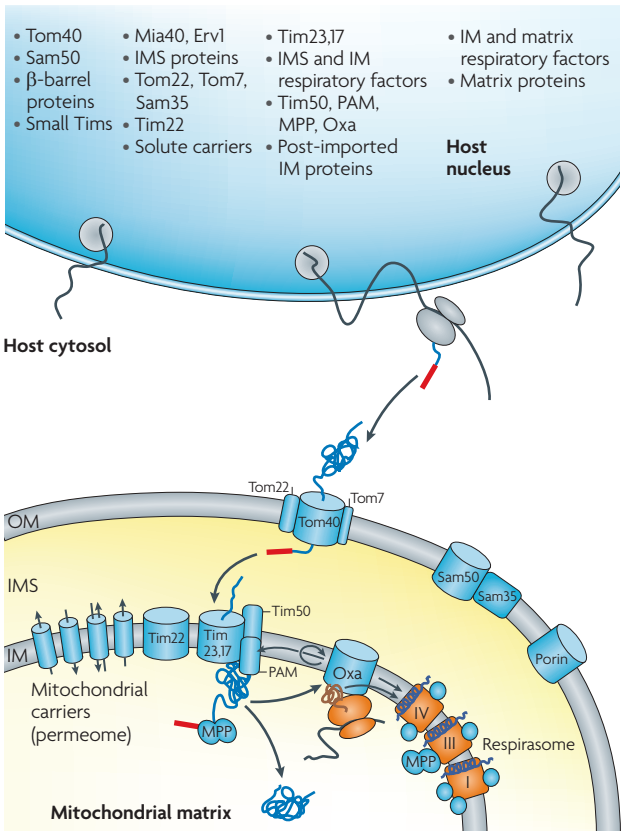
a Mitochondrial organellogenesis



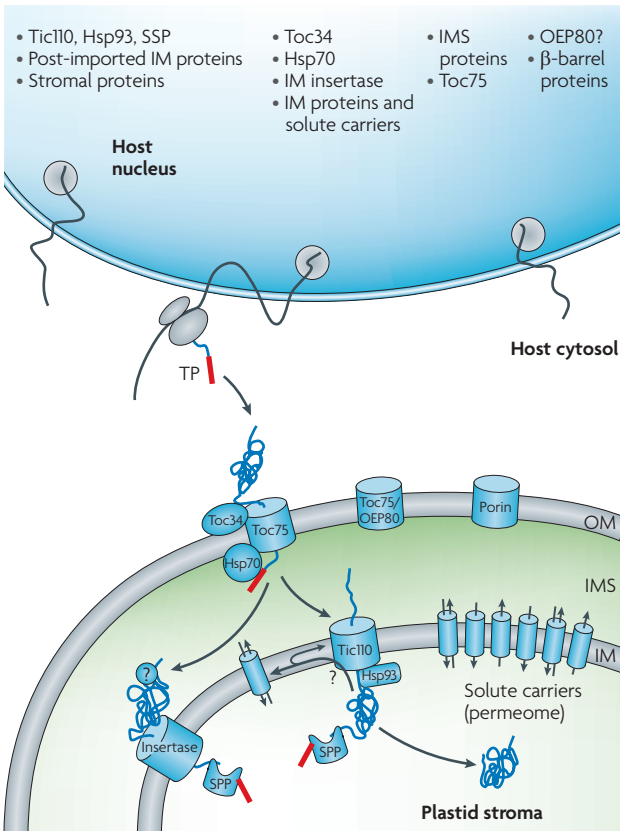
b Plastid organellogenesis



c Mitochondrial organellogenesis



d Plastid organellogenesis



◀ Figure 2 | **Later events in mitochondrial and plastid organellogenesis.** Molecular factors are either encoded in the nucleus (blue) or in the alphaproteobacterial (yellow) or cyanobacterial (green) chromosome. Question marks indicate currently unknown components. **a** | The Tim22 insertase was introduced in the alphaproteobacterial inner membrane (IM) to catalyse integration of the mitochondrial carriers and its own insertion. Small Tim proteins bridged the transit of sorting substrates from Tom40 to the Tim22 insertase. Tim23 and Tim17 were substrates of Tim22 inserted in the IM to constitute another protein pore. This pore evolved to integrate STMD proteins containing presequences (red) into the IM as respiratory factors (for example, supernumerary subunits, small blue circles) to control the biogenesis of the alphaproteobacterial respiratory complexes I, III and IV (the respirasome). An alphaproteobacterial cytosolic peptidase was recruited to the respirasome complex III to cleave the presequences of Tim23 substrates. Other soluble respiratory factors (for example, supernumerary subunits) could potentially have been attached to the sector of the respirasome exposed to the intermembrane space (IMS). **b** | A putative nuclear-encoded insertase evolved to integrate nuclear-encoded solute carriers in the cyanobacterial IM. Its function was supported by hypothetical nuclear-encoded components (chaperones) that were previously established in the IMS. At this stage the transit peptides (TPs, red) could have evolved either to facilitate the translocation of substrates across the outer membrane (OM) by binding an IMS Hsp70 motor, or to help the integration of solute carriers by the IM insertase. Toc34 was established as a receptor for TP-containing proteins. **c** | The putative ancestral protein-sorting system in mitochondria[4,24]. The addition of the presequence translocase-associated motor (PAM) module to the Tim23 insertase allowed the translocation of proteins into the mitochondrial matrix. Among them were new respiratory factors acting in the respirasome sector exposed to the matrix. Matrix processing peptidase (MPP) was derived from the alphaproteobacterial peptidase when the corresponding gene was established in the nucleus. Pools of MPP became associated with the repirasome complex III. Oxa was established as a mitoribosome recruiting factor to promote the cotranslational insertion of respirasome subunits encoded in the mitochondrial genome. In addition, Oxa provided a pathway for post-import insertion in the IM of nuclear-encoded proteins (for example, Oxa, PAM subunits and some respirasome components)[22]. **d** | The putative ancestral protein-sorting system in plastids[21]. The Tic110 translocase evolved in connection with the Hsp93 chaperone motor, making possible the translocation of proteins into the plastid stroma. Among them was the nuclear-encoded stromal processing peptidase (SPP). The Tic complex might have provided a pathway for post-import insertion of IM proteins such as solute carriers and Tic110.

core 1 and core 2 subunits of complex III of the mitochondrial respiratory chain[54–58]. In many fungi (including *Neurospora crassa*) the β-subunit of the MPP is the same as the core 1 subunit[55]. In addition, in all green algae and plants, core 1 and core 2 polypeptides are the only catalytic peptidases involved in the maturation of Tim23 presequences[54,56]. Also, the activity of presequence peptidase has been experimentally demonstrated for the core 1 and core 2 subunits of complex III in *Bos taurus*[57]. We propose that the tendency for MPP activity to associate physically with complex III (FIG. 2c), in many instances conserved during evolution, reflects an early recruitment to the respiratory chain of the original alphaproteobacterial peptidase[58] (FIG. 2a) to allow maturation of the typical substrates of Tim23; that is, subunits of the respirasome and adjunct biogenetic factors. Intriguingly, in modern yeast mitochondria, Tim23 is physically associated with complexes III and IV of the respirasome[42,46].

Protein import into the matrix evolved later. We propose that selective pressure to enhance the function of Tim23 as a central adduct for biogenetic components of the respiratory chain triggered the next transition in mitochondrial organellogenesis: the evolution of protein translocation across the IM. This was made possible by the addition of the presequence translocase-associated motor (PAM) module to the Tim23 insertase (FIG. 2c); this module was built around a recruited heat shock protein, Hsp70, located in the matrix of mitochondria[22,46,59]. Evolution of the PAM module promoted the nuclear establishment of an array of new components, the products of which could be imported across the IM in an ATP-driven reaction to act in the sector of the respiratory chain exposed to the inner organelle (FIG. 2c). These included several supernumerary subunits[51–53], proteins involved in the biosynthesis of cofactors[53], catalytic subunits that are inserted in the IM from inside the organelle or that are exposed to the matrix[52], and biogenetic factors[53]. In addition, the insertase Oxa, which was redesigned in the nuclear genome by addition of a C-terminal ribosome tethering domain, allowed the host to gain control over cotranslational insertion into the organelle IM of respirasome subunits encoded in the mitochondrial genome[60] (FIG. 2c), which removed their biogenetic dependence on endogenous prokaryotic Sec and Tat translocons. Thus, the refinement of control over the alphaproteobacterial biogenetic activities reached the sorting pathways in the organelle interior.

Plastid inner membrane before the stroma

Transit peptides before stromal targeting. What insights does our outsiders' perspective provide about protein translocation systems in the IMS and IM of plastids in light of the shortage of experimental data in this arena? Current views largely support the notion that, by default, plastid translocons manage the targeting of proteins containing a cleavable N-terminal extension (a transit peptide, TP) to the organelle interior[20,40,61]. Despite this consensus, proteins destined to the plastid IMS and IM surprisingly also contain TPs[40,61,62], whereas most OM proteins do not[29,61]. We propose that this reflects the fact that TPs evolved in the context of protein targeting to IMS or the IM of the primitive plastid, after protein sorting to the OM was established but before import into the organelle interior was feasible. Because of the similarities between TPs and Tim23 presequences[61,63], it is possible that primitive TPs might have originated from mitochondrial presequences by exon shuffling and were selected owing to their high affinity for Hsp70 chaperones[64]. It could be that one Hsp70 protein was established in the IMS after the advent of Toc75 to stimulate the full translocation of proteins across the OM by binding to their TPs[64] (FIG. 2b). Interestingly, the TP is the first region of the imported protein intermediate that emerges from the Toc75 pore. Alternatively, it is conceivable that a binding specificity for TPs existed in the *trans* site of a putative IM insertase (see below) that potentially directed the precursors to partially enter the stroma, thereby facilitating the correct positioning of the membrane anchor signal inside the IM insertase (FIG. 2b). This is analogous to how N-terminal extensions function in the Tim23 insertase[22,42]. Therefore, in the context of both IMS and IM sorting, TPs might have initially been established not as topogenic signals but as topogenic devices (BOX 3).

An insertase for permeome reprogramming? The mechanism of integration of some IM proteins in plastids, including solute transporters, proceeds from the IMS and involves their insertion in a putative IM channel followed by lateral release of the cargo in the lipid bilayer[40,65,66]. By analogy to mitochondrial Tim23 IM substrates, the presence of a membrane anchor signal following a TP is sufficient to integrate the targeted proteins in the IM of plastids[40,65,66]. The core component of

Box 3 | **Evolving the semantics of protein sorting**

Organellogenesis depended on reliable mechanisms to ensure that nuclear-encoded proteins would reach their suborganellar destination. It is traditionally suggested that a 'code' system evolved on the basis of topogenic signals in the transported proteins to 'inform' their address in the organelle[70]. The best example of such a sorting tag is the N-terminal presequence, purportedly a 'master signal' for protein import into the mitochondrial and plastid interior. However, the contradiction of this notion is that the evolution of such an appended code system is of little use without the simultaneous establishment of receptor properties to decode the topogenic signals.

Our perspective, gained from studying mitochondrial evolution, suggests a more parsimonious scenario in which presequences initially evolved as structures added to single transmembrane domain (STMD) proteins to increase their rate of insertion into the inner membrane (IM). To enhance their substrate specificity, Tom40 and the Tim23 insertase evolved efficient gating that was dependent on the recognition of the presequences[22,43], which were the first region of the substrates to insert into the protein pores. In addition, Tim23 evolved an excellent mechanism for lateral release of STMD proteins in the IM by 'sensing' the adjustment of the hydrophobic transmembrane domain in the core of the insertase embedded in the lipid bilayer[22,42,45,46]. Therefore, it was initially the affinity of protein pores for structural tracts common to all their substrates that made sense of the presequences and membrane anchor regions. The reverse logic of this semantic process is that once presequences and membrane anchors were endowed with a 'meaning' they became 'signals' — that is, a signal to target a protein to the Tim23 pore and to insert this protein in the IM. This specificity was further enhanced by addition of subunits to the pores — these subunits would act as protein receptors for the topogenic signals (for example, Tom22 (REF. 74)) (FIG. 1c). Over time, the co-evolution of pores, receptors and the topogenic signals strengthened the specificity and fine-tuned the fidelity of the protein-sorting process.

The idea that topogenic signals evolved in the context of distinctive structural features that pre-existed in the sorted proteins is further illustrated in the plastid. In this organelle, transit peptides (TPs) might also have been added to intermembrane space (IMS) and IM proteins as transient catalytic helpers for the import reaction. But, in contrast to mitochondria, in plastids it was the TPs that evolved their specificity to fit the receptor properties that pre-existed in the Omp85-derived pore (that is, Toc75). This included adaptation of a phenylalanine residue in the N-terminus of TPs that is a key residue in the topogenic signal of β-barrel proteins recognized by Omp85 pores[27]. In this way, TPs evolved the status of a signal for Toc75-mediated import towards the IMS and the IM. It is not by chance that the further insertion of proteins in the IM of plastids seems to have recapitulated the recognition of the hydrophobic transmembrane anchor segment in the insertion substrate[65,66]. An intriguing possibility is that the lack of a suitable membrane anchor signal might have been selected during evolution of both mitochondria and plastids as the authentic signal to direct imported proteins into the organelle interior.

the membrane anchor signal is likewise an α-helical hydrophobic domain[65,66]. All of these characteristics are compatible with a protein-sorting system that pre-dates the targeting to the interior of the organelle and probably results in direct insertion of preproteins into the IM rather than a 'stop of the transfer' to the plastid stroma. Interestingly, a recent study that analysed the import of different plastid IM proteins seems to indicate that all tested precursors uniformly cross the OM through the Toc translocon, but from their import behaviour it seems they diverge afterwards in two different groups, probably reflecting their requirement in two distinct IM pores[40]. Because one of the tested proteins, PIC1, has recently been shown to be inserted from the stroma after crossing the IM[67], it is conceivable that the set of proteins that includes PIC1 uses the Tic translocon. Alternatively, the remaining group of proteins might converge to a second, yet to be identified insertase. We postulate that this might be the standard IM insertase that evolved during organellogenesis to control the reprogramming of the cyanobacterial IM permeome with eukaryote-derived solute transporters (FIG. 2b).

Protein import into the stroma evolved later. Speculation about an early IM insertase in plastids raises the question of whether this component facilitated the implementation of the subsequent step in plastid organellogenesis: the evolution of protein translocation across the IM by the Tic translocon. The current model is that the primitive Tic translocon was composed solely of a protein-conducting pore, the eukaryotic-derived Tic110, coupled to the stromal chaperone Hsp93 (REFS 21,68) (FIG. 2d). In light of our outsiders' perspective, we propose two different scenarios that link the putative pre-existing IM insertase to the subsequent evolution of the primitive Tic translocon.

In the first scenario, the IM insertase might have been a biogenetic factor of Tic110 (for example, by catalysing its integration in the IM). Thereby, the IM insertase that was implemented early could have promoted the establishment of Tic110 in an analogous manner to the evolution of Tim23 and Tim17 in the mitochondrial IM that might have followed the establishment of the Tim22 insertase. Alternatively, it is conceivable that Tic110 originally evolved as the pore of the IM insertase. Later, by interacting with different subunits, Tic110 acquired the ability to translocate proteins into the plastid stroma. In this way, two different Tic110 subcomplexes would account for protein insertion in the IM and stromal translocation in modern plastids[40]. It is possible that protein import to the stroma was ensured by interaction of Tic110 with the chaperone motor Hsp93 (REF. 21) (FIG. 2d). This would recapitulate the paradigm described above for mitochondrial matrix import involving the recruitment of the inner organelle Hsp70 to the IM insertase Tim23 (REF. 46). Interestingly, the Tic translocon seems to have been further adapted to the insertion of IM proteins that are integrated in the lipid bilayer after crossing the IM (FIG. 2d), because post-import integration of these proteins in modern plastids apparently depends on the Tic components Tic110, Tic40 and the Hsp93 chaperone[67,69,70]. As in mitochondrial evolution, once the plastid host gained access to the organelle interior it assumed command over internal sorting pathways by incorporating new molecular components that took over the functions of the endogenous Sec and Tat translocons.

Conclusions

We suggest that molecular analysis of the protein-sorting functions in modern-day plastids will provide the best test for the ideas put forward here. A fundamental issue to be clarified is the extent of correlation between the mitochondrial and plastid protein-sorting systems. This correspondence would be substantiated in our hypothesis by meeting three key predictions. First, the assembly of OM β-barrel proteins in plastids might involve the Omp85-derived components Toc75 or OEP80. Second, molecular factors in the IMS, analogous to the small Tim chaperones, might support the function of components in the

OM and IM. Finally, a biogenetic or functional connection between the putative IM insertase and the Tic complex might reflect the evolutionary transition from IM insertion to translocation across the IM. Furthermore, a detailed understanding of protein sorting to the plastid membranes will provide the opportunity to re-evaluate the prevalent notion of primacy of protein sorting to the interior of the organelles. This insiders' perspective traces back to the signal hypothesis that originally stated that, by default, protein channels in the endoplasmic reticulum membrane translocate proteins containing signal peptides to the compartment lumen[71]. Our hypothetical model of prokaryote-to-organelle conversion points instead to the primacy of the membranes as targets for mitochondrial and plastid evolution, and warrants a re-evaluation of the traditional paradigm for protein translocation. We postulate that import of proteins to the interior of organelles could only arise as an add-on to previously established OM, IMS and IM sorting systems. These developments opened up a new phase of organelle evolution that is marked by extensive nuclear control over the internal organellar processes, including organelle division, metabolism, genome replication and gene expression.

Here we postulate that access by nuclear-encoded factors to the outermost compartments of the Gram-negative endosymbionts laid down the path for both mitochondrial and plastid establishment. Would such a constrained outside-to-inside trajectory be a paradigm for endosymbiont-to-organelle conversion? We suggest that the ideal test of our hypothesis will be provided by the characterization of endosymbiont-derived organelles that potentially have recapitulated the outsiders' trajectory for their establishment. The best current candidates are secondary plastids derived from algal endosymbionts that were engulfed by protist hosts — that is, chromalveolates, chlorarachniophytes and euglenophytes[2]. Additional examples may be provided by *Carsonella ruddii* in the phloem sap-feeding psyllids[72] and the cyanobacterial-derived cyanelle (that is, plastid) in the photosynthetic amoeba *Paulinella chromatophora*[73]. Given their biogenetic dependence on their respective eukaryotic hosts, these Gram-negative-bacteria-derived intracellular compartments seem to be genuine organelles. Notably, as for mitochondria and plastids, in all the examples described above the eukaryotic host is the dominant player in prokaryote-to-organelle conversion. This suggests that nuclear control over endosymbiont fate is an intrinsic aspect of organellogenesis. This idea stands in clear contrast to the alternative view that the nucleus arose because of an equal or dominant role of the alphaproteobacterial endosymbiont[7,8]. The sort of nuclear dominance over endosymbiont-to-organelle conversion highlighted here suggests that the host of the alphaproteobacterial mitochondrial forerunner was either a bona fide eukaryote[3,6] or, alternatively, a prokaryote in the early stages of eukaryogenesis, but whose fate was not dominated by the alphaproteobacterial endosymbiont.

Jeferson Gross and Debashish Bhattacharya are at the Department of Biology and the Roy J. Carver Center for Comparative Genomics, University of Iowa, 446 Biology Building, Iowa City, Iowa 52242, USA.

Correspondence to D.B.
e-mail: debashi-bhattacharya@uiowa.edu

doi:10.1038/nrg2610
Published online 9 June 2009

Glossary

β-Barrel proteins
A class of membrane proteins composed of antiparallel β-strands that form barrel-type pores. Porins are typical β-barrel pores in the outer membrane of Gram-negative bacteria.

Chaperones
These are molecular components that fold, unfold, stabilize or escort the transit of protein substrates. In addition, chaperones such as the 70 kDa heat shock protein (Hsp70) and Hsp93 hydrolyse ATP to provide the energy for protein import across organelle membranes.

Endosymbiotic gene transfer
The migration and fixation of endosymbiont genes in the nuclear genome of the host.

Muller's ratchet
Describes the progressive irreversible accumulation of deleterious mutations in asexual populations. Muller's ratchet explains the genomic and physiological degeneration that is usually observed among obligatory endosymbionts.

Membrane anchor signal
This is a topogenic signal used to anchor a membrane protein in the lipid bilayer. The core of a membrane anchor signal is usually the first hydrophobic α-helix that is C-terminal to the presequence.

Presequence
This is a transient topogenic signal appended to the N-terminus of a sorted protein that is cleaved from the mature protein.

Presequence translocase-associated motor
(PAM). This is a module of the Tim23 complex associated with Hsp70 that provides the energy for translocation of proteins across the mitochondrial inner membrane. In yeast, PAM is composed of the structural platform Tim44, the co-chaperones Pam14, Pam16 and Pam18, and the nucleotide exchange factor Mge1.

Single transmembrane domain
(STMD). α-Helical STMDs are found in membrane proteins, and fold into a hydrophobic, helical structure spanning the lipid bilayer.

Sorting and assembly machinery
(SAM). The SAM complex inserts and assembles β-barrel proteins in the OM of mitochondria. In yeast, it is comprised of the protein pore Sam50 and the peripheral subunits Sam35 and Sam37.

Sorting substrate
This is a protein that is relocated by the catalytic action of a translocase or an insertase. A sorting substrate is also referred to as a precursor or a preprotein.

Tic
(Translocon at the inner envelope membrane of chloroplasts). In higher plants, the Tic complex is composed of the protein-conducting channel Tic110, the putative protein pores Tic20 and Tic21, the intermembrane space protein Tic22, the chaperone Hsp93 and its co-chaperone Toc40. Tic32, Tic55 and Tic62 are regulatory subunits.

Tim22 insertase
An insertase is a molecular machine usually consisting of a receptor and a protein pore that recognizes specific protein substrates and catalyses their insertion into the lipid bilayer of a membrane. The Tim22 insertase is specialized in the insertion of the mitochondrial carriers into the inner membrane. In yeast, it is comprised of the Tim22 protein pore and the subunits Tim54 and Tim18.

Tim23
(Translocase of the mitochondrial inner membrane). In yeast, the Tim23 and Tim17 subunits constitute the protein-conducting pore of the Tim23 complex that, in combination with Tim50 and Tim21, acts as an insertase for single transmembrane domain proteins. The addition of the PAM module confers a translocase function for the Tim23 complex.

Toc
(Translocon at the outer envelope membrane of chloroplasts). In higher plants, the Toc complex is composed of the protein-conducting pore Toc75, the receptors Toc34 and Toc159, and the accessory subunits Toc12 and Toc64.

Tom
(Translocase of the mitochondrial outer membrane). In yeast, the Tom complex is formed by the Tom40 protein-conducting pore, the receptors Tom22, Tom20 and Tom70, and the structural subunits Tom5, Tom6 and Tom7

Topogenic signal
This is a segment of amino acids in a sorted protein that provides information about its final location and conformation (topology) in a cellular compartment.

Translocon
A molecular machine usually consisting of a receptor and a protein pore that recognizes specific protein substrates and catalyses their movement across a membrane. 'Translocase' is a general term to describe an enzyme that facilitates the movement of a molecule, usually across a membrane.

1. Bhattacharya, D., Archibald, J. M., Weber, A. P. & Reyes-Prieto, A. How do endosymbionts become organelles? Understanding early events in plastid evolution. *Bioessays* **29**, 1239–1246 (2007).
2. Reyes-Prieto, A., Weber, A. P. & Bhattacharya, D. The origin and establishment of the plastid in algae and plants. *Annu. Rev. Genet.* **41**, 147–168 (2007).
3. de Duve, C. The origin of eukaryotes: a reappraisal. *Nature Rev. Genet.* **8**, 395–403 (2007).
4. Dolezal, P., Likic, V., Tachezy, J. & Lithgow, T. Evolution of the molecular machines for protein import into mitochondria. *Science* **313**, 314–318 (2006).
5. Dyall, S. D., Brown, M. T. & Johnson, P. J. Ancient invasions: from endosymbionts to organelles. *Science* **304**, 253–257 (2004).
6. Cavalier-Smith, T. Origin of mitochondria by intracellular enslavement of a photosynthetic purple bacterium. *Proc. Biol. Sci.* **273**, 1943–1952 (2006).
7. Martin, W. & Muller, M. The hydrogen hypothesis for the first eukaryote. *Nature* **392**, 37–41 (1998).
8. Martin, W. & Koonin, E. V. Introns and the origin of nucleus–cytosol compartmentalization. *Nature* **440**, 41–45 (2006).
9. Yoon, H. S., Hackett, J. D., Ciniglia, C., Pinto, G. & Bhattacharya, D. A molecular timeline for the origin of photosynthetic eukaryotes. *Mol. Biol. Evol.* **21**, 809–818 (2004).
10. Martin, W. & Herrmann, R. G. Gene transfer from organelles to the nucleus: how much, what happens, and why? *Plant Physiol.* **118**, 9–17 (1998).
11. Kurland, C. G. & Andersson, S. G. Origin and evolution of the mitochondrial proteome. *Microbiol. Mol. Biol. Rev.* **64**, 786–820 (2000).
12. Andersson, S. G., Karlberg, O., Canback, B. & Kurland, C. G. On the origin of mitochondria: a genomics perspective. *Philos. Trans. R. Soc. Lond. B* **358**, 165–177 (2003); discussion 177–179.
13. Richly, E. & Leister, D. An improved prediction of chloroplast proteins reveals diversities and commonalities in the chloroplast proteomes of *Arabidopsis* and rice. *Gene* **329**, 11–16 (2004).
14. Panigrahi, A. K. *et al.* A comprehensive analysis of *Trypanosoma brucei* mitochondrial proteome. *Proteomics* **9**, 434–450 (2009).
15. Moran, N. A. Accelerated evolution and Muller's rachet in endosymbiotic bacteria. *Proc. Natl Acad. Sci. USA* **93**, 2873–2878 (1996).
16. Reyes-Prieto, A., Hackett, J. D., Soares, M. B., Bonaldo, M. F. & Bhattacharya, D. Cyanobacterial contribution to algal nuclear genomes is primarily limited to plastid functions. *Curr. Biol.* **16**, 2320–2325 (2006).
17. Tyra, H. M., Linka, M., Weber, A. P. & Bhattacharya, D. Host origin of plastid solute transporters in the first photosynthetic eukaryotes. *Genome Biol.* **8**, R212 (2007).
18. Moustafa, A., Reyes-Prieto, A. & Bhattacharya, D. Chlamydiae has contributed at least 55 genes to Plantae with predominantly plastid functions. *PLoS ONE* **3**, e2205 (2008).
19. Reumann, S. & Keegstra, K. The endosymbiotic origin of the protein import machinery of chloroplastic envelope membranes. *Trends Plant Sci.* **4**, 302–307 (1999).
20. Reumann, S., Inoue, K. & Keegstra, K. Evolution of the general protein import pathway of plastids (review). *Mol. Membr. Biol.* **22**, 73–86 (2005).
21. Gross, J. & Bhattacharya, D. Revaluating the evolution of the Toc and Tic protein translocons. *Trends Plant Sci.* **14**, 13–20 (2009).
22. Neupert, W. & Herrmann, J. M. Translocation of proteins into mitochondria. *Annu. Rev. Biochem.* **76**, 723–749 (2007).
23. Kutik, S., Guiard, B., Meyer, H. E., Wiedemann, N. & Pfanner, N. Cooperation of translocase complexes in mitochondrial protein import. *J. Cell Biol.* **179**, 585–591 (2007).
24. Lister, R., Hulett, J. M., Lithgow, T. & Whelan, J. Protein import into mitochondria: origins and functions today (review). *Mol. Membr. Biol.* **22**, 87–100 (2005).
25. Herrmann, J. M. Converting bacteria to organelles: evolution of mitochondrial protein sorting. *Trends Microbiol.* **11**, 74–79 (2003).
26. Ruiz, N., Kahne, D. & Silhavy, T. J. Advances in understanding bacterial outer-membrane biogenesis. *Nature Rev. Microbiol.* **4**, 57–66 (2006).
27. Wunder, T., Martin, R., Loffelhardt, W., Schleiff, E. & Steiner, J. M. The invariant phenylalanine of precursor proteins discloses the importance of Omp85 for protein translocation into cyanelles. *BMC Evol. Biol.* **7**, 236 (2007).
28. Inoue, K. & Keegstra, K. A polyglycine stretch is necessary for proper targeting of the protein translocation channel precursor to the outer envelope membrane of chloroplasts. *Plant J.* **34**, 661–669 (2003).
29. Inoue, K. & Potter, D. The chloroplastic protein translocation channel Toc75 and its paralog OEP80 represent two distinct protein families and are targeted to the chloroplastic outer envelope by different mechanisms. *Plant J.* **39**, 354–365 (2004).
30. Patel, R., Hsu, S. C., Bedard, J., Inoue, K. & Jarvis, P. The Omp85-related chloroplast outer envelope protein OEP80 is essential for viability in *Arabidopsis*. *Plant Physiol.* **148**, 235–245 (2008).
31. Walther, D. M., Papic, D., Bos, M. P., Tommassen, J. & Rapaport, D. Signals in bacterial β-barrel proteins are functional in eukaryotic cells for targeting to and assembly in mitochondria. *Proc. Natl Acad. Sci. USA* **106**, 2531–2356 (2009).
32. Duy, D., Soll, J. & Philippar, K. Solute channels of the outer membrane: from bacteria to chloroplasts. *Biol. Chem.* **388**, 879–889 (2007).
33. Gentle, I. E. *et al.* Conserved motifs reveal details of ancestry and structure in the small TIM chaperones of the mitochondrial intermembrane space. *Mol. Biol. Evol.* **24**, 1149–1160 (2007).
34. Alcock, F. H. *et al.* Conserved substrate binding by chaperones in the bacterial periplasm and the mitochondrial intermembrane space. *Biochem. J.* **409**, 377–387 (2008).
35. Beverly, K. N., Sawaya, M. R., Schmid, E. & Koehler, C. M. The Tim8–Tim13 complex has multiple substrate binding sites and binds cooperatively to Tim23. *J. Mol. Biol.* **382**, 1144–1156 (2008).
36. Allen, J. W., Ferguson, S. J. & Ginger, M. L. Distinctive biochemistry in the trypanosome mitochondrial intermembrane space suggests a model for stepwise evolution of the MIA pathway for import of cysteine-rich proteins. *FEBS Lett.* **582**, 2817–2825 (2008).
37. Weber, A. P. & Fischer, K. Making the connections — the crucial role of metabolite transporters at the interface between chloroplast and cytosol. *FEBS Lett.* **581**, 2215–2222 (2007).
38. Kunji, E. R. The role and structure of mitochondrial carriers. *FEBS Lett.* **564**, 239–244 (2004).
39. Schneider, A. *et al.* An *Arabidopsis thaliana* knock-out mutant of the chloroplast triose phosphate/phosphate translocator is severely compromised only when starch synthesis, but not starch mobilisation is abolished. *Plant J.* **32**, 685–699 (2002).
40. Firlej-Kwoka, E., Strittmatter, P., Soll, J. & Bolter, B. Import of preproteins into the chloroplast inner envelope membrane. *Plant Mol. Biol.* **68**, 505–519 (2008).
41. Rassow, J., Dekker, P. J., van Wilpe, S., Meijer, M. & Soll, J. The preprotein translocase of the mitochondrial inner membrane: function and evolution. *J. Mol. Biol.* **286**, 105–120 (1999).
42. van der Laan, M. *et al.* A role for Tim21 in membrane-potential-dependent preprotein sorting in mitochondria. *Curr. Biol.* **16**, 2271–2276 (2006).
43. van der Laan, M. *et al.* Motor-free mitochondrial presequence translocase drives membrane integration of preproteins. *Nature Cell Biol.* **9**, 1152–1159 (2007).
44. Popov-Celeketic, D., Mapa, K., Neupert, W. & Mokranjac, D. Active remodelling of the TIM23 complex during translocation of preproteins into mitochondria. *EMBO J.* **27**, 1469–1480 (2008).
45. Gruhler, A. *et al.* N-terminal hydrophobic sorting signals of preproteins confer mitochondrial hsp70 independence for import into mitochondria. *J. Biol. Chem.* **272**, 17410–17415 (1997).
46. Wiedemann, N., van der Laan, M., Hutu, D. P., Rehling, P. & Pfanner, N. Sorting switch of mitochondrial presequence translocase involves coupling of motor module to respiratory chain. *J. Cell Biol.* **179**, 1115–1122 (2007).
47. Acin-Perez, R., Fernandez-Silva, P., Peleato, M. L., Perez-Martos, A. & Enriquez, J. A. Respiratory active mitochondrial supercomplexes. *Mol. Cell* **32**, 529–539 (2008).
48. Richter, O. M. & Ludwig, B. Cytochrome c oxidase — structure, function, and physiology of a redox-driven molecular machine. *Rev. Physiol. Biochem. Pharmacol.* **147**, 47–74 (2003).
49. Cardol, P. *et al.* The mitochondrial oxidative phosphorylation proteome of *Chlamydomonas reinhardtii* deduced from the Genome Sequencing Project. *Plant Physiol.* **137**, 447–459 (2005).
50. Brandt, U. *et al.* Structure–function relationships in mitochondrial complex I of the strictly aerobic yeast *Yarrowia lipolytica*. *Biochem. Soc. Trans.* **33**, 840–844 (2005).
51. Zara, V., Conte, L. & Trumpower, B. L. Biogenesis of the yeast cytochrome bc_1 complex. *Biochim. Biophys. Acta* **1793**, 89–96 (2009).
52. Lazarou, M., Thorburn, D. R., Ryan, M. T. & McKenzie, M. Assembly of mitochondrial complex I and defects in disease. *Biochim. Biophys. Acta* **1793**, 78–88 (2009).
53. Fontanesi, F., Soto, I. C., Horn, D. & Barrientos, A. Assembly of mitochondrial cytochrome *c*-oxidase, a complicated and highly regulated cellular process. *Am. J. Physiol. Cell Physiol.* **291**, C1129–C1147 (2006).
54. Howell, K. A. *et al.* Oxygen initiation of respiration and mitochondrial biogenesis in rice. *J. Biol. Chem.* **282**, 15619–15631 (2007).
55. Schulte, U. *et al.* A family of mitochondrial proteins involved in bioenergetics and biogenesis. *Nature* **339**, 147–149 (1989).
56. Glaser, E. & Dessi, P. Integration of the mitochondrial-processing peptidase into the cytochrome bc_1 complex in plants. *J. Bioenerg. Biomembr.* **31**, 259–274 (1999).
57. Deng, K., Shenoy, S. K., Tso, S. C., Yu, L. & Yu, C. A. Reconstitution of mitochondrial processing peptidase from the core proteins (subunits I and II) of bovine heart mitochondrial cytochrome bc_1 complex. *J. Biol. Chem.* **276**, 6499–6505 (2001).
58. Brown, M. T. *et al.* A functionally divergent hydrogenosomal peptidase with protomitochondrial ancestry. *Mol. Microbiol.* **64**, 1154–1163 (2007).
59. Frazier, A. E. *et al.* Pam16 has an essential role in the mitochondrial protein import motor. *Nature Struct. Mol. Biol.* **11**, 226–233 (2004).
60. Bonnefoy, N., Fiumera, H. L., Dujardin, G. & Fox, T. D. Roles of Oxa1-related inner-membrane translocases in assembly of respiratory chain complexes. *Biochim. Biophys. Acta* **1793**, 60–70 (2009).
61. Jarvis, P. Targeting of nucleus-encoded proteins to chloroplasts in plants. *New Phytol.* **179**, 257–285 (2008).
62. Vojta, L., Soll, J. & Bolter, B. Protein transport in chloroplasts — targeting to the intermembrane space. *FEBS J.* **274**, 5043–5054 (2007).
63. Brink, S. *et al.* Preproteins of chloroplast envelope inner membrane contain targeting information for receptor-dependent import into fungal mitochondria. *J. Biol. Chem.* **269**, 16478–16485 (1994).
64. Jackson-Constan, D., Akita, M. & Keegstra, K. Molecular chaperones involved in chloroplast protein import. *Biochim. Biophys. Acta* **1541**, 102–113 (2001).
65. Brink, S., Fischer, K., Klosgen, R. B. & Flugge, U. I. Sorting of nuclear-encoded chloroplast membrane proteins to the envelope and the thylakoid membrane. *J. Biol. Chem.* **270**, 20808–20815 (1995).
66. Knight, J. S. & Gray, J. C. The N-terminal hydrophobic region of the mature phosphate translocator is sufficient for targeting to the chloroplast inner envelope membrane. *Plant Cell* **7**, 1421–1432 (1995).
67. Chiu, C. C. & Li, H. M. Tic40 is important for reinsertion of proteins from the chloroplast stroma into the inner membrane. *Plant J.* **56**, 793–801 (2008).
68. Balsera, M. *et al.* Characterization of Tic110, a channel-forming protein at the inner envelope membrane of chloroplasts, unveils a response to Ca^{2+} and a stromal regulatory disulfide bridge. *J. Biol. Chem.* **284**, 2603–2616 (2009).
69. Inaba, T. *et al.* Arabidopsis tic110 is essential for the assembly and function of the protein import machinery of plastids. *Plant Cell* **17**, 1482–1496 (2005).
70. Vojta, L., Soll, J. & Bolter, B. Requirements for a conservative protein translocation pathway in chloroplasts. *FEBS Lett.* **581**, 2621–2624 (2007).
71. Blobel, G. Intracellular protein topogenesis. *Proc. Natl Acad. Sci. USA* **77**, 1496–1500 (1980).
72. Tamames, J. *et al.* The frontier between cell and organelle: genome analysis of *Candidatus* Carsonella ruddii. *BMC Evol. Biol.* **7**, 181 (2007).
73. Yoon, H. S., Reyes-Prieto, A., Melkonian, M. & Bhattacharya, D. Minimal plastid genome evolution in the *Paulinella* endosymbiont. *Curr. Biol.* **16**, R670–672 (2006).
74. Yamano, K. *et al.* Tom20 and Tom22 share the common signal recognition pathway in mitochondrial protein import. *J. Biol. Chem.* **283**, 3799–3807 (2008).

75. Kutik, S. *et al.* Dissecting membrane insertion of mitochondrial β-barrel proteins. *Cell* **132**, 1011–1024 (2008).
76. Zientz, E., Dandekar, T. & Gross, R. Metabolic interdependence of obligate intracellular bacteria and their insect hosts. *Microbiol. Mol. Biol. Rev.* **68**, 745–770 (2004).
77. Janssen, M. J., Koorengevel, M. C., de Kruijff, B. & de Kroon, A. I. Transbilayer movement of phosphatidylcholine in the mitochondrial outer membrane of *Saccharomyces cerevisiae* is rapid and bidirectional. *Biochim. Biophys. Acta* **1421**, 64–76 (1999).
78. de Azevedo-Martins, A. C., Frossard, M. L., de Souza, W., Einicker-Lamas, M. & Motta, M. C. Phosphatidylcholine synthesis in *Crithidia deanei*: the influence of the endosymbiont. *FEMS Microbiol. Lett.* **275**, 229–236 (2007).
79. Benning, C. A role for lipid trafficking in chloroplast biogenesis. *Prog. Lipid Res.* **47**, 381–389 (2008).
80. Lee, D. H., Severin, K., Yokobayashi, Y. & Ghadiri, M. R. Emergence of symbiosis in peptide self-replication through a hypercyclic network. *Nature* **390**, 591–594 (1997).
81. Chacinska, A. *et al.* Mitochondrial biogenesis, switching the sorting pathway of the intermembrane space receptor Mia40. *J. Biol. Chem.* **283**, 29723–29729 (2008).

Acknowledgements
D.B. acknowledges support from the National Science Foundation and the National Institutes of Health. The authors also thank W. Lanier (Iowa) for a critical reading of the manuscript.

FURTHER INFORMATION

Jeferson Gross's homepage: http://www.biology.uiowa.edu/debweb/LabMsites/Jeferson_Gross.php
Debashish Bhattacharya's homepage: http://cyanophora.biology.uiowa.edu/home
NCBI organelle genome resources: http://www.ncbi.nlm.nih.gov/genomes/genlist.cgi?taxid=2759&type=4&name=Eukaryotae%20Organelles

SUPPLEMENTARY INFORMATION

See online article: S1 (table)

ALL LINKS ARE ACTIVE IN THE ONLINE PDF