



Mini-review

Origin and evolution of metabolic sub-cellular compartmentalization in eukaryotes



Toni Gabaldón ^{a, b, c, *}, Alexandros A. Pittis ^{a, b}

^a Bioinformatics and Genomics Programme, Centre for Genomic Regulation (CRG), Dr. Aiguader, 88, 08003 Barcelona, Spain

^b Universitat Pompeu Fabra (UPF), 08003 Barcelona, Spain

^c Institució Catalana de Recerca i Estudis Avançats (ICREA), Pg. Lluís Companys 23, 08010 Barcelona, Spain

ARTICLE INFO

Article history:

Received 6 February 2015

Accepted 25 March 2015

Available online 11 April 2015

Keywords:

Evolution

Eukaryotes

Compartments

Organelles

ABSTRACT

A high level of subcellular compartmentalization is a hallmark of eukaryotic cells. This intricate internal organization was present already in the common ancestor of all extant eukaryotes, and the determination of the origins and early evolution of the different organelles remains largely elusive. Organellar proteomes are determined through regulated pathways that target proteins produced in the cytosol to their final subcellular destinations. This internal sorting of proteins can vary across different physiological conditions, cell types and lineages. Evolutionary retargeting – the alteration of a subcellular localization of a protein in the course of evolution – has been rampant in eukaryotes and involves any possible combination of organelles. This fact adds another layer of difficulty to the reconstruction of the origins and evolution of organelles. In this review we discuss current themes in relation to the origin and evolution of organellar proteomes. Throughout the text, a special focus is set on the evolution of mitochondrial and peroxisomal proteomes, which are two organelles for which extensive proteomic and evolutionary studies have been performed.

© 2015 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

1. Introduction

The diversity of cellular life on earth can be divided into three main domains: bacteria, archaea and eukaryotes [1]. Eukaryotes are defined by the presence of an intracellular membrane-bound compartment that harbors the chromosomes, in contrast to bacteria and archaea – commonly referred to as prokaryotes [2]– in which chromosomes are contained in the cytosol. In addition to the nucleus, even the simplest eukaryotic cell displays an intricate network of intracellular membranes that define compartments, also known as organelles, in which specific biological processes take place. Some prokaryotes do display specialized sub-cellular domains such as magnetosomes in magnetotactic bacteria [3] or photosynthetic thylakoids in cyanobacteria [4]. In this regard, members of the *Planctomycetes*, *Verrucomicrobiae*, and *Chlamydiae* (PVC) bacterial superphylum are exceptional in that they exhibit an extensive intracellular membranous organization. However, detailed analysis has determined that these structures result from expansions of the plasma membrane, and that there is no physical

separation of transcription and translation [5]. Hence, in terms of the complexity of their subcellular organization, there is a great divide separating the simplest eukaryotic cell and the most complex cells among prokaryotes. In the absence of organisms with an intermediate complexity, we are left with the difficult question of how the intricate subcellular compartmentalization in eukaryotes originated in the first place. Solving this puzzle remains one of the fundamental challenges of evolutionary biology [6]. This difficulty notwithstanding, several hypotheses have been put forward that explain the origin of different eukaryotic organelles. Assessing the validity of these hypotheses is extremely difficult, because they refer to very ancient events whose footprints have certainly been extensively altered and eroded in part, if not completely, through time. Consequently, we can only expect to obtain indirect evidence from these events. A sensible approach in these circumstances is to assess the plausibility of the proposed models in the context of our current biochemical, cellular and evolutionary knowledge, and contrast their predictions with the growing landscape of available genomic data.

It is not our intention to provide a comprehensive account of proposed evolutionary paths to eukaryotic compartmentalization in this review. Rather, we will just briefly mention the main hypotheses to later move on to discussing general concepts and issues

* Corresponding author. Bioinformatics and Genomics Programme, Centre for Genomic Regulation (CRG), Dr. Aiguader, 88, 08003 Barcelona, Spain.

E-mail address: tgabaldon@crg.es (T. Gabaldón).

regarding the origin and subsequent evolution of cellular compartments. In this regard a focus of the review is set on the evolutionary retargeting of proteins and pathways to other compartments, a theme which is not often included in reviews dealing with eukaryogenesis. In the first section a brief survey of the main proposed evolutionary scenarios for the principal types of organelles that can be found in a typical eukaryotic cell is presented. We discuss them in terms of their consistency with existing phylogenetic and biochemical data. Once a sub-cellular compartmentalization is in place, this provides an opportunity to separate metabolic processes that would otherwise interfere, but it also poses important requirements for the regulation of protein traffic. How proteins are distributed to the different compartments and how this process is regulated is discussed in the second section of this review. Furthermore, given a system to specifically sort a repertoire of proteins among a set of different compartments, there is a multitude of plausible arrangements that would result in different metabolic capacities. This, in turn, provides a new dimension for exploring the phenotypic space by altering the subcellular localization of proteins. Such re-wiring of the subcellular proteome is a well known mechanism that leads to functional differences between cells across different conditions or tissues, and recent evidence is showing that this rewiring also occurs during evolution. In the third section of the review we discuss evidence showing that subcellular retargeting has been common during the early and recent evolution of eukaryotes. Throughout the article we will focus on concrete examples related to the evolution of the mitochondrial and peroxisomal proteomes, two of the organelles whose evolutionary trajectories have been most extensively studied.

2. The origin of cellular compartments: endogenous vs exogenous routes

Current hypotheses on the evolutionary origin of eukaryotic organelles can be broadly divided into two main groups according to whether the organelle was formed from pre-existing structures within the cell (endogenous or autogenous) or from the interaction of different cells that established a symbiotic relationship (exogenous or xenogenous). It is important to note that these two different scenarios also confront different paradigms, in which either ramification (endogenous) or merging (exogenous) represents the main driving force of evolution. As we will see, these evolutionary paradigms, that do not necessarily exclude each other, have been confronted several times. Exogenous hypotheses take usually the form of organelles derived from ancestral (endo)symbiotic interactions between different cellular organisms. Such hypotheses are inherently attractive because they readily provide a mechanism for a preformed compartment with particular metabolic properties. Ideas of an endosymbiotic origin for plastids and mitochondria have been around since the early times of cellular microscopy [7,8], but these were only broadly discussed later, when Lynn Margulis proposed the serial endosymbiotic theory [9]. This theory extended the idea of exogenous origins beyond plastids and mitochondria, proposing bacterial origins for the basal bodies of flagella. These endosymbiotic ideas initially received heavy opposition, but they were broadly accepted—at least for mitochondria and plastids—by the 1980s, after compelling evidence was found that these two organelles harbored their own bacterial-like genome and translation machinery [10,11]. These early confirmatory results fueled the search for exogenous origins for other organelles, and many hypotheses were put forward that explain the origin of organelles in terms of symbiotic interactions among prokaryotes (Fig. 1 and Table 1). As mentioned above, Lynn Margulis argued that eukaryotic flagella descend from bacterial structures,

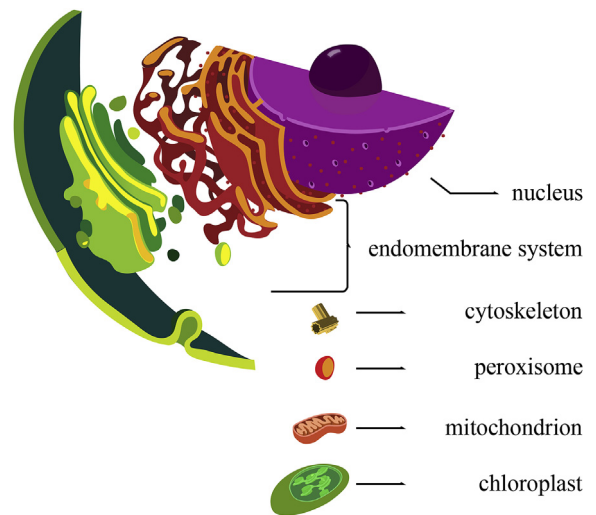


Fig. 1. Eukaryotic compartments. Schematic representation of a generalized eukaryotic cell showing the main compartments. For simplicity, and because there is agreement on their common origin, we have grouped into the single category “endomembrane system” the endoplasmic reticulum, Golgi apparatus, lysosomes, vacuoles, vesicles, and endosomes. Main hypotheses regarding the origins of these organelles are reviewed in Table 1. Figure has been modified from wikipedia.

Table 1

Main hypotheses regarding the exogenous or endogenous origin of the different organelles (See Fig. 1).

Organelle	Autogenous origin	Symbiotic/Endosymbiotic origin
Nucleus	[2,13,15,40,41]	[42–44]
Cytoskeleton	[14,45]	[46–48]
Endomembrane system	[15,49]	[42,50]
Mitochondria	[51]	[9,45,46,52]
Plastids		[46,53,54]
Peroxisomes	[16,19,20]	[17,55]

through the fusion of a spirochete cell with an archaeon, also giving rise to the cytoskeleton [9]. The proposal was made on the basis of morphological similarities, but also on the observation of spirochete attachments to the plasma membrane of some anaerobic protists, which were used as motility organelles [12]. Even for the nucleus, exogenous origins have been suggested. In most of these models the nucleus would descend from an archaeal cell that established endosymbiosis within a bacterial host (for a review see Ref. [13]). A bacterial endosymbiotic origin was also proposed for peroxisomes soon after it was realized that new peroxisomes can form out of pre-existing ones through growth and division (see below). However in most cases the quest for a definitive evidence of endosymbiotic origin of the type that is available for mitochondria and plastids has been elusive. For instance, in the case of cytoskeleton, and flagellar bodies in particular, subsequent molecular comparative analysis failed to show a signal of spirochete ancestry among the corresponding proteins. Instead many of them proved to be eukaryotic-specific or only distantly related to prokaryotic groups other than spirochetes, as in the case of actin and tubulin, the major structural components of cytoskeleton [14]. Similarly, accumulating data has not definitely supported so far any particular scenario of a supposed endosymbiotic origin for the nucleus, and there is no current consensus as to this point. In the case of peroxisomes molecular data similarly did not reveal any particularly strong footprint of the peroxisomal proteome for any bacterial group. All in all, given that mitochondria and plastids are the only DNA bearing organelles other than the nucleus, there has been no clear evidence pointing to an endosymbiotic origin of other DNA-

lacking eukaryotic organelles. However, this cannot be taken as definitive evidence of endogenous origin, as there are examples of complete genome loss in mitochondria, as it is the case for mitosomes and hydrogenosomes. This makes evolutionary models for the origin of organelles difficult to test.

The absence of confirmatory evidence for many proposed exogenous model cooled down symbiotic enthusiasm, and prompted researchers to envision alternative, endogenous scenarios (Fig. 1 and Table 1), with no fewer mechanistic and conceptual difficulties. As the endosymbiotic origin of eukaryotic flagella was never widely accepted by the scientific community, models of the gradual development of eukaryotic cytoskeleton in a stepwise manner were developed. As in all aspects of eukaryogenesis, we have no evidence of how such a process could have evolved, as intermediate steps are missing from the records, but prevailing models consider cytoskeleton to have evolved from an archaeal-like host and before the last eukaryotic common ancestor [14]. For the endomembrane system, including the nuclear envelope, a step-wise development through pre-existing membranes has been proposed. In some models this is presented as a direct response to stimulus from the mitochondrial endosymbiont, as in the case of the hydrogen hypothesis [15]. Similarly, several lines of evidence coming from molecular data and cell biological observations reinforced autogenous models for the origin of the peroxisome. Indeed, the peroxisome is one paradigmatic case in which preferred hypotheses for the origin of the organelles have shifted from exogenous to endogenous scenarios [16]. Based on the observation that peroxisomes, similarly to mitochondria and plastids, multiply by division of preexisting ones, and that they possess a translocation machinery for the import of proteins synthesized in the cytosol, Christian de Duve postulated an endosymbiotic origin of this organelle [17]. In his view, peroxisomes descended from a prokaryotic endosymbiont able to cope with oxidative damage, which provided an essential advantage to the eukaryotic ancestor at a time when molecular oxygen was rising in the earth's atmosphere. The absence of an organellar genome and the diversity of metabolic pathways in peroxisomes from different lineages suggested a very ancestral endosymbiosis –probably predating that of mitochondria– of a metabolically complex ancestor. This complex ancestor would have later diversified by differential loss of pathways in the different lineages, while always keeping a central role in the management of reactive oxygen species. This view was later challenged by a growing disparity of peroxisomal metabolic pathways as new lineages of microbial eukaryotes were analyzed, and by the finding of ontogenic and evolutionary connections between these organelles and the endomembrane system. In particular, the protein translocation machinery, which is fundamentally different from those involved in protein import to mitochondria and peroxisomes, was found to be homologous to the Endoplasmic Reticulum Associated Decay (ERAD) system [18–20]. This evolutionary connection fitted recent experimental data showing the budding of new peroxisomes from the Endoplasmic Reticulum in certain yeast mutants [21]. Altogether, these data supported an emergent alternative scenario in which peroxisomes originated endogenously, being yet other type of off-shots of the endomembrane system. Although this view is now broadly accepted, the specific scenario in which peroxisomes arose, and what were the evolutionary forces that drove the process remain under debate [22–24].

3. Directing the traffic: protein sorting mechanisms

With the exception of the relatively few that are encoded by the mitochondrial and plastid genomes, proteins are synthesized in cytosolic ribosomes and are directed post-translationally to their final destination by means of an intricate system of sorting routes.

At its core, a subcellular sorting mechanism has two main components. First, a targeting signal, which is contained within the polypeptide chain or in the folded structure of the protein. Secondly, specific soluble receptors, which recognize these signals and direct the proteins to a specific membrane-associated translocation machinery that drives the insertion of the protein in the lumen or membrane of a particular compartment. The specificity of this sorting mechanism is defined by a so-called “sorting code” in which different signals are recognized by specific receptors and translocators. Similarly to the epigenetic code [25], the sorting code is not static and posttranslational modifications and conformational changes can activate or deactivate a given signal. In addition, the targeting of the same protein to more than one different compartment, perhaps under different physiological conditions, is common, indicating that different targeting signals can coexist.

The diversity and fundamental differences of some of the targeting signals and translocation systems indicate that at least several types of sorting mechanisms have originated independently. Conversely, some sorting systems have more or less obvious similarities at the sequence and the mechanistic level that indicate that the sorting pathways have also diversified by duplication and co-option of their components. One such example is the peroxisomal import machinery, which shows remarkable homology and mechanistic parallels to the ERAD system [19]. Interestingly, this system was later co-opted in alga and apicomplexan parasites, which have obtained plastids by secondary endocytosis, i.e. endocytosis of an alga containing a primary plastid. In these organisms ERAD homologs form the symbiont-specific ERAD-like machinery (SELMA) system, a pre-protein translocation machinery that directs proteins to the secondary plastids in these organisms [26]. It has been proposed that the efficiency of ancestral targeting systems enabled massive subcellular retargeting across compartments [27]. Indeed it is tempting to envision that, in a context with soft-targeting rules and pervasive dual targeting of proteins, the retargeting of multiple inter-related enzymes such as those acting on the same pathway may have been more plausible. Conversely, this model implies that the relocation of proteins would have become more and more unlikely as the sorting mechanisms increased their specificity.

4. Diversity and evolutionary variation of sub-cellular proteomes

Organelles are dynamic entities. Morphological differences between the same organelles in different species, cell-types, or physiological conditions have been noted [28,29]. These differences were also observed in the form of the presence of different metabolic properties, which indicated that different enzymatic functions were targeted to the same organelles in different cell-types, organisms or under different physiological conditions. For instance, in mammals, mitochondria from different types of skeletal muscle fibers have significant morphological and physiological differences. Indeed, the aerobic type I muscle fibers contain many mitochondria and rely on oxidative phosphorylation for the production of ATP and can oxidize glycolytic products but also fatty acids and ketones. In contrast mitochondria of Type IIa fibers are numerous but smaller than those in Type I, and can only oxidize glycolytic products. Finally, Type IIb fibers contain few mitochondria, harbor high levels of glycogen and mainly produce ATP via anaerobic glycolysis. Similarly large differences have been found in almost any other type of eukaryotic organelle. More recently, the possibility to target the localization of specific proteins and perform proteomic analysis of specific compartments allowed examination of the actual protein repertoire of a given organelle across different species, tissues, or conditions [28]. These analyses have revealed that organellar

proteomes not only differ from cell type to cell type but also that they respond dynamically to several types of perturbations. In this review we will limit our discussion to differences across organellar composition in different lineages because they reveal the evolutionary dynamics of organellar proteomes and provide evidence for the existence of evolutionary retargeting, which will be the focus of the next section.

Considering that an organelle's proteome is variable across tissues and conditions, the comparison of organellar proteomes across different species is a complex task and goes well beyond efficiently recognising orthologs across the compared species. An absolute description of the proteome of an organelle in a given species would ideally include all proteins targeted therein in any set of physiological conditions and tissues. Alternatively one can limit the comparison to physiological conditions that are largely comparable across the considered species, which is not an easy task for species that are relatively distant. In practice, most comprehensive subcellular catalogs have been built for several species by compiling available data from numerous experiments performed under different conditions with different techniques and in various labs. Although still an approximation, these catalogs are useful to provide an idea of how variable an organellar proteome is across lineages. Large subcellular proteomic datasets have been compiled for mitochondria in human, rat, mouse, and baker's yeast [30]. By comparing such broad datasets in human and yeast, it has been estimated that the mitochondria of these two species share only about 42–46% of their total proteome [31]. This low degree of overlap is remarkable considering that the two species belong to the same broad clade of eukaryotes (opisthokonts). The common mitochondrial proteomic set includes many core mitochondrial pathways, such as most of the oxidative phosphorylation complexes, the assembly of Fe–S clusters, most of the proteins of the mitochondrial carrier family, and the protein synthesis and import machineries. Examples of proteins that are present in mitochondria from human but absent from those in yeast include NADH:ubiquinone oxidoreductase (Complex I), fatty acid beta-oxidation, steroid biogenesis, and the apoptotic Bcl2-family signaling pathway. Conversely, trehalose synthesis, glycerone-P metabolism, and starch and sucrose degradation appear to be specific for yeast mitochondria. While a fraction of the pathways that are present differentially in the two species correspond to differential loss of genes in the corresponding lineages (e.g. loss of Complex I in *Saccharomyces*), a significant fraction do have orthologs in the other species, but the encoded proteins and pathways have different subcellular destinations. This is the case of beta-oxidation of long-chain fatty acids, which has a mitochondrial localization in mammals but it is peroxisomal in yeast.

5. Proteins on the move: re-targeting as an evolutionary playground

The observed diversity of the proteome composition in organelles across different lineages points to the existence of evolutionary retargeting events, i.e. the alteration of the sub-cellular localization of a protein throughout evolution. How often this happens and what are the evolutionary forces that play a role in this process is still a matter of study. As explained above, the targeting signals are generally encoded by short stretches of amino acid sequences. Hence, any given protein is not many mutations away from acquiring or losing a given targeting signal. A concrete example that serves to illustrate this point is provided by the evolution of the subcellular localization of the alanine:glyoxylate aminotransferase (AGT), which has altered its subcellular localization in numerous occasions during the evolution of mammals in parallel to changes in dietary habits. This enzyme tends to be

localized to mitochondria in carnivores (including insectivores), to peroxisomes in herbivores, and usually displays a dual localization to both organelles in omnivores [32]. AGT catalyzes the detoxification of the intermediary metabolite glyoxylate to glycine, thereby preventing its oxidation to oxalate. Oxalate accumulation is toxic because its calcium salt is poorly soluble, leading to the formation of harmful crystals in the kidney and urinary tract. Obviously, for AGT to efficiently prevent calcium oxalate crystals it has to be localized precisely at the site of glyoxylate synthesis. The best evidence of a deleterious effect of an incorrect subcellular localization of AGT is found in humans, where a single point mutation causing mistargeting of AGT to mitochondria, instead of peroxisomes, is potentially lethal in patients suffering from kidney stone disease primary hyperoxaluria type 1 (PH1) [33]. Notably the site of glyoxylate production is likely to be different in herbivores and carnivores, because the main precursor for glyoxylate is either glycolate (metabolized in peroxisomes) or hydroxyproline (metabolized in mitochondria), respectively [34]. Accordingly, the relative localization of the enzyme correlates with diet in different species, and has changed at least three times during the evolution of mammals in lineages where the diet has shifted: bats, primates, and the order carnivora (Fig. 2). Most significantly, among members of the carnivora, the shift from mitochondrial localization to peroxisomal localization in the lineage leading to the Giant panda, a strict herbivore, has been associated to an excess of non-synonymous changes in the region coding for the mitochondrial targeting signal. These changes, in turn, resulted in a reduced efficiency of mitochondrial targeting and an increase of peroxisomal localization in the Giant panda as compared to its omnivorous close relatives. This is a prime example of how natural selection can drive the relocalization of proteins to different compartments. In other documented examples a gene duplication has predated the retargeting event, in which one of the two resulting paralogous genes is targeted to a new compartment [19]. In this regard, it is interesting to note that pairs of paralogs are more likely to display a different subcellular localization as compared to orthologs and that subcellular localization is the ontology category that shows stronger differences in the behavior of orthologs and paralogs [35].

To gain a more comprehensive insight into the extent of retargeting between the different organelles we compared experimentally-based annotations of homologous proteins from model organisms. For this we defined protein families by combining gene trees of homologous proteins found in PhylomeDB [36] and tracked how often experimentally-based annotations of different subcellular locations, as documented by the Gene Ontology annotations at Uniprot [37], were co-occurring (see Fig. 3 legend for further details). The results of this rough survey show that, in a minimal estimate, 32% of eukaryotic protein families had at least two homologous proteins targeted to different compartments, which is evidence for at least one event of retargeting. Thus, at this evolutionary scale, re-localization of proteins to different compartments has been rampant. The particular arrangements of the observed co-occurrences among different compartments (Fig. 3) suggest that any possible retargeting between two compartments is possible, but some combinations are more common than others. This may reflect underlying topological, evolutionary and functional links among the organelles. Strikingly, almost two thirds of the peroxisomal proteome corresponds to proteins with homologs re-targeted to or from mitochondria, more than half to or from nucleus, whereas fewer of them have homologs localized to the other compartments. Nonetheless, we observe a significant overlap between the nuclear envelope, the plasma membrane and the endomembrane system, pointing to a pool of protein families alternatively targeted to these organelles or dually localized among them. Future studies that exploit a larger body of subcellular

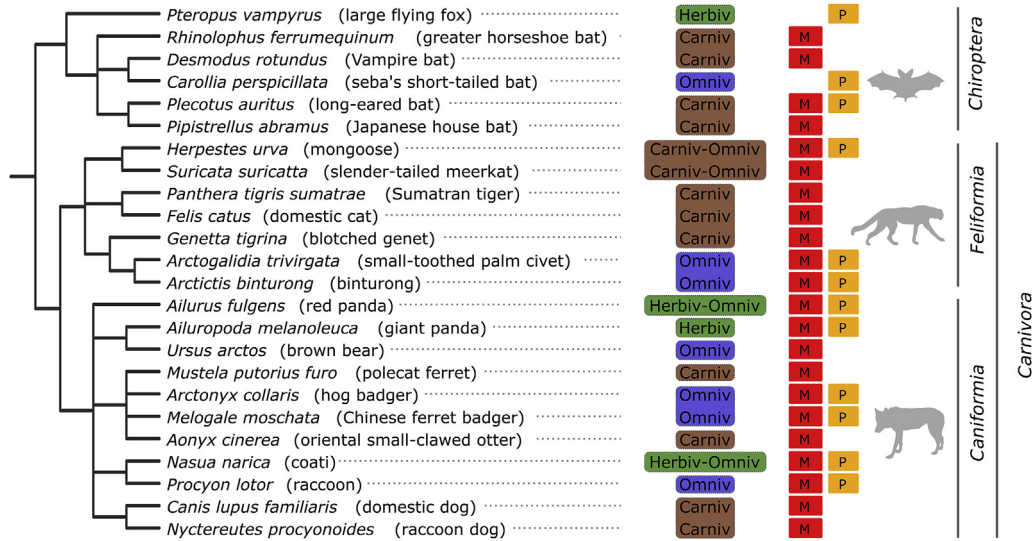


Fig. 2. Subcellular retargeting of AGT in relation to diet. NCBI taxonomy tree of 24 species indicating the relation between the subcellular distribution of alanine:glyoxylate aminotransferase (AGT) in liver cells and their diet. M: mitochondrial, P: peroxisomal. Carniv: carnivorous, Omniv: omnivorous, Herbiv: herbivorous, Carniv–Omniv: mainly carnivorous, but some plant material eaten, Herbiv–Omniv: mainly herbivorous, but some animal material eaten. Adapted from Ref. [38]. ETE was used for tree visualization [39].

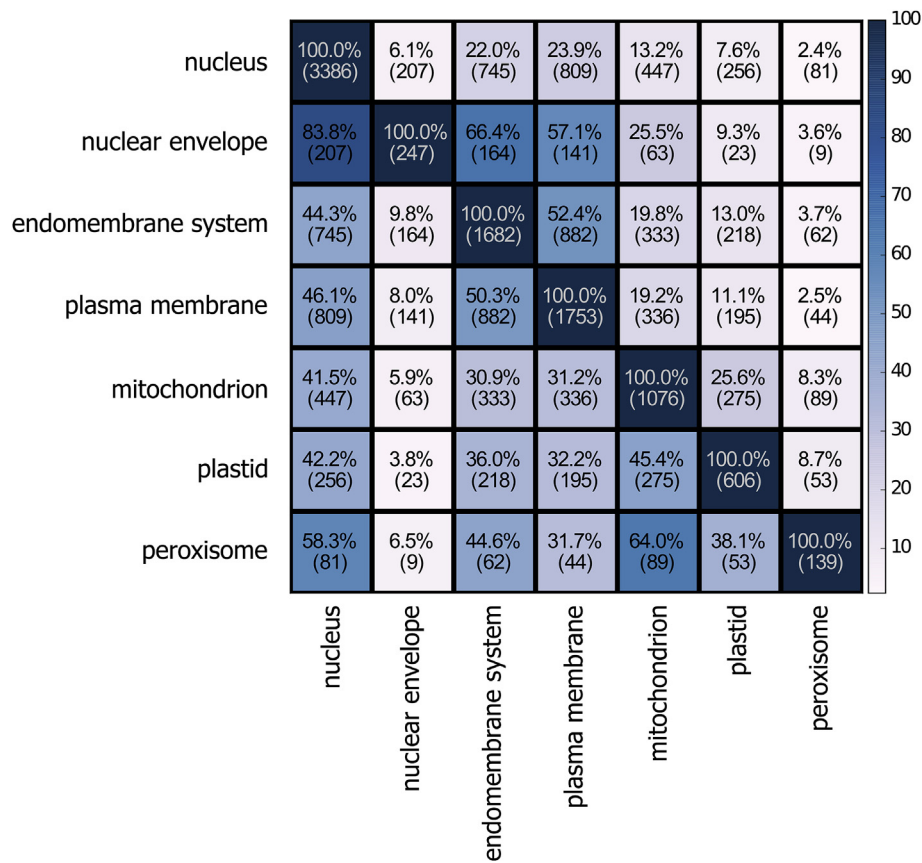


Fig. 3. Degree of retargeting between the different eukaryotic compartments. The frequency of co-occurrence of different organellar location within the same protein family is indicated as a heatmap. The count (in parentheses) indicates the number of proteins from the same family targeted to both organelles (alternatively in different species, or dually-targeted in the same species). Percentages refer to the fraction of protein families targeted to both organelles compared to the overall size of the proteome of a given organelle (in the diagonal), in a row-wise manner. Notably, almost 2/3 of the peroxisomal proteome comprises protein families with members that localize in mitochondria. Retargeting was evaluated by looking at the co-occurrence of the various localizations within the same protein families. We defined the families by combining phylogenetic trees from two phylomes containing mostly model organisms (PhylomeIDs 500 and 502) and fusing trees with overlapping sequences based on UPGMA clustering [36]. Subsequently, for any possible pair of two compartments, we counted the number of times that two distinct experimentally-based annotations of subcellular localization co-occurred in the same family and plotted in a heat-map the counts and the percentage over the total number of families in that organelle. For a term to be considered for a protein family, at least 2 annotated proteins with the given term were required. Only annotation terms based on actual experiments were considered.

localization and genomic data will help establishing the dynamics of evolutionary retargeting in more detail.

6. Concluding remarks

The subcellular compartmentalization of eukaryotic cells is complex and offers the possibility of selectively isolating metabolic processes and biochemical reactions that may interfere with each other. The targeting of proteins to the different locations is tightly regulated in a time and spacial manner. Accumulating evidence show that many proteins have been retargeted through evolution and that organellar proteomes have been shaped significantly. This process has likely played a major role in the adaptation of organisms to different environments, a role that we are now just starting to understand. On the other hand evolutionary retargeting blurs the evolutionary signal of organellar ancestry and complicates its reconstruction, adding to the other challenges of elucidating early eukaryotic evolution. The study of evolutionary retargeting of proteins is in its infancy. We now have a growing number of interesting examples but lack comprehensive studies that expand complete proteomes across various taxa. The scarcity of reliable annotations for subcellular localizations in non-model organisms prevents obtaining accurate figures, but early analyses suggest that evolutionary retargeting has been –and currently is– rampant, providing a new tinkering tool for evolution. Improvements in large-scale determination of subcellular retargeting, and growing genomic and proteomic datasets for a larger diversity of eukaryotes will certainly help us in understanding this important evolutionary process.

Conflict of interest

None.

Acknowledgments

TG group research is funded in part by a grant from the Spanish ministry of Economy and Competitiveness (BIO2012-37161), a Grant from the Qatar National Research Fund grant (NPRP 5-298-3-086), and a grant from the European Research Council under the European Union's Seventh Framework Programme (FP/2007-2013)/ERC (Grant Agreement n. ERC-2012-StG-310325).

References

- [1] C.R. Woese, O. Kandler, M.L. Wheelis, Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya, *Proc. Natl. Acad. Sci. U. S. A.* 87 (12) (1990 Jun) 4576–4579.
- [2] W. Martin, E.V. Koonin, A positive definition of prokaryotes, *Nature* 442 (7105) (2006 Aug 24) 868.
- [3] D.A. Bazylinski, R.B. Frankel, Magnetosome formation in prokaryotes, *Nat. Rev. Microbiol.* 2 (3) (2004 Mar) 217–230.
- [4] J. Nickelsen, B. Rengstl, A. Stengel, M. Schottkowski, J. Soll, E. Ankele, Biogenesis of the cyanobacterial thylakoid membrane system—an update, *FEMS Microbiol. Lett.* 315 (1) (2011 Feb) 1–5.
- [5] R. Santarella-Mellwig, S. Pruggnaller, N. Roos, I.W. Mattaj, D.P. Devos, Three-dimensional reconstruction of bacteria with a complex endomembrane system, *PLoS Biol.* 11 (5) (2013) e1001565.
- [6] E.V. Koonin, Darwinian evolution in the light of genomics, *Nucleic Acids Res.* 37 (4) (2009 Mar) 1011–1034.
- [7] C. Mereschkowsky, Über Natur und Ursprung der Chromatophoren im Pflanzenreiche, *Biol. Cent.* 25 (1905) 593–604.
- [8] I.E. Wallin, Symbiogenesis and the Origin of Species, Bailliere, Tindall and Cox, London, 1927, p. 171.
- [9] L. Sagan, On the origin of mitosing cells, *J. Theor. Biol.* 14 (3) (1967 Mar) 255–274.
- [10] M.W. Gray, W.F. Doolittle, Has the endosymbiont hypothesis been proven? *Microbiol. Rev.* 46 (1) (1982 Mar) 1–42.
- [11] M.W. Gray, G. Burger, B.F. Lang, Mitochondrial evolution, *Science* 283 (5407) (1999 Mar 5) 1476–1481.
- [12] A.M. Wier, L. Sacchi, M.F. Dolan, C. Bandi, J. Macalister, L. Margulis, Spirochete attachment ultrastructure: implications for the origin and evolution of cilia, *Biol. Bull.* 218 (1) (2010 Feb) 25–35.
- [13] W. Martin, Archaeobacteria (Archaea) and the origin of the eukaryotic nucleus, *Curr. Opin. Microbiol.* 8 (6) (2005 Dec) 630–637.
- [14] B. Wickstead, K. Gull, The evolution of the cytoskeleton, *J. Cell. Biol.* 194 (4) (2011 Aug 22) 513–525.
- [15] W. Martin, M. Müller, The hydrogen hypothesis for the first eukaryote, *Nature* 392 (6671) (1998 Mar 5) 37–41.
- [16] T. Gabaldón, Peroxisome diversity and evolution, *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 365 (1541) (2010 Mar 12) 765–773.
- [17] C. De Duve, Evolution of the Peroxisome, *Ann. N. Y. Acad. Sci.* 168 (2) (1969) 369–381.
- [18] T. Gabaldón, S. Capella-Gutiérrez, Lack of phylogenetic support for a supposed actinobacterial origin of peroxisomes, *Gene* 465 (1–2) (2010 Oct 1) 61–65.
- [19] T. Gabaldón, B. Snel, F. van Zimmeren, W. Hemrika, H. Tabak, M.A. Huynen, Origin and evolution of the peroxisomal proteome, *Biol. Direct* 1 (2006) 8.
- [20] A. Schlüter, S. Fourcade, R. Ripp, J.L. Mandel, O. Poch, A. Pujol, The evolutionary origin of peroxisomes: an ER-peroxisome connection, *Mol. Biol. Evol.* 23 (4) (2006 Apr) 838–845.
- [21] D. Hoepfner, D. Schildknegt, I. Braakman, P. Philippsen, H.F. Tabak, Contribution of the endoplasmic reticulum to peroxisome formation, *Cell* 122 (1) (2005 Jul 15) 85–95.
- [22] T. Gabaldón, A metabolic scenario for the evolutionary origin of peroxisomes from the endomembranous system, *Cell. Mol. Life Sci. CMLS* 71 (13) (2014 Jul) 2373–2376.
- [23] T. Gabaldón, Evolutionary considerations on the origin of peroxisomes from the endoplasmic reticulum, and their relationships with mitochondria, *Cell. Mol. Life Sci. CMLS* 71 (13) (2014 Jul) 2379–2382.
- [24] D. Speijer, Oxygen radicals shaping evolution: why fatty acid catabolism leads to peroxisomes while neurons do without it: FADH₂/NADH flux ratios determining mitochondrial radical formation were crucial for the eukaryotic invention of peroxisomes and catabolic tissue differentiation, *BioEssays News Rev. Mol. Cell. Dev. Biol.* 33 (2) (2011 Feb) 88–94.
- [25] B.M. Turner, Defining an epigenetic code, *Nat. Cell. Biol.* 9 (1) (2007 Jan) 2–6.
- [26] K. Bolte, N. Gruenheit, G. Felsner, M.S. Sommer, U.-G. Maier, F. Hempel, Making new out of old: recycling and modification of an ancient protein translocation system during eukaryotic evolution. Mechanistic comparison and phylogenetic analysis of ERAD, SELMA and the peroxisomal importomer, *BioEssays News Rev. Mol. Cell. Dev. Biol.* 33 (5) (2011 May) 368–376.
- [27] W. Martin, Evolutionary origins of metabolic compartmentalization in eukaryotes, *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 365 (1541) (2010 Mar 12) 847–855.
- [28] J.S. Andersen, M. Mann, Organellar proteomics: turning inventories into insights, *EMBO Rep.* 7 (9) (2006 Sep) 874–879.
- [29] F. Forner, L.J. Foster, S. Campanaro, G. Valle, M. Mann, Quantitative proteomic comparison of rat mitochondria from muscle, heart, and liver, *Mol. Cell. Proteomics MCP* 5 (4) (2006 Apr) 608–619.
- [30] S.W. Taylor, E. Fahy, S.S. Ghosh, Global organellar proteomics, *Trends Biotechnol.* 21 (2) (2003 Feb) 82–88.
- [31] T. Gabaldón, M.A. Huynen, From endosymbiont to host-controlled organelle: the hijacking of mitochondrial protein synthesis and metabolism, *PLoS Comput. Biol.* 3 (11) (2007 Nov) e219.
- [32] C.J. Danpure, Variable peroxisomal and mitochondrial targeting of alanine: glyoxylate aminotransferase in mammalian evolution and disease, *BioEssays News Rev. Mol. Cell. Dev. Biol.* 19 (4) (1997 Apr) 317–326.
- [33] C.J. Danpure, The molecular basis of alanine: glyoxylate aminotransferase mistargeting: the most common single cause of primary hyperoxaluria type 1, *J. Nephrol.* 11 (Suppl. 1) (1998 Apr) 8–12.
- [34] T. Takayama, K. Fujita, K. Suzuki, M. Sakaguchi, M. Fujie, E. Nagai, et al., Control of oxalate formation from L-hydroxyproline in liver mitochondria, *J. Am. Soc. Nephrol. JASN* 14 (4) (2003 Apr) 939–946.
- [35] T. Gabaldón, E.V. Koonin, Functional and evolutionary implications of gene orthology, *Nat. Rev. Genet.* 14 (5) (2013 May) 360–366.
- [36] J. Huerta-Cepas, S. Capella-Gutiérrez, L.P. Pryszcz, M. Marcet-Houben, T. Gabaldón, PhylomeDB v4: zooming into the plurality of evolutionary histories of a genome, *Nucleic Acids Res.* 42 (Database issue) (2014 Jan) D897–D902.
- [37] R.P. Huntley, T. Sawford, P. Mutowo-Muullenet, A. Shypitsyna, C. Bonilla, M.J. Martin, et al., The GOA database: gene ontology annotation updates for 2015, *Nucleic Acids Res.* 43 (Database issue) (2014 Nov 6) D1057–D1063.
- [38] G.M. Birdsey, J. Lewin, A.A. Cunningham, M.W. Bruford, C.J. Danpure, Differential enzyme targeting as an evolutionary adaptation to herbivory in carnivora, *Mol. Biol. Evol.* 21 (4) (2004 Apr) 632–646.
- [39] J. Huerta-Cepas, J. Dopazo, T. Gabaldón, ETE: a python environment for tree exploration, *BMC Bioinforma.* 11 (2010) 24.
- [40] T. Cavalier-Smith, Origin of the cell nucleus, *BioEssays News Rev. Mol. Cell. Dev. Biol.* 9 (2–3) (1988 Sep) 72–78.
- [41] T. Cavalier-Smith, Origin of the cell nucleus, mitosis and sex: roles of intracellular coevolution, *Biol. Direct.* 5 (2010) 7.
- [42] R.S. Gupta, G.B. Golding, The origin of the eukaryotic cell, *Trends Biochem. Sci.* 21 (5) (1996 May) 166–171.
- [43] T. Horiike, K. Hamada, S. Kanaya, T. Shinozawa, Origin of eukaryotic cell nuclei by symbiosis of Archaea in bacteria is revealed by homology-hit analysis, *Nat. Cell. Biol.* 3 (2) (2001 Feb) 210–214.

- [44] J.A. Lake, M.C. Rivera, Was the nucleus the first endosymbiont? *Proc. Natl. Acad. Sci. U. S. A.* 91 (8) (1994 Apr 12) 2880–2881.
- [45] L. Margulis, Serial Endosymbiosis Theory, in: *Symbiosis in Cell Evolution: Microbial Communities in the Archean and Proterozoic Eons*, 1993, pp. 1–18.
- [46] M.W. Gray, The endosymbiont hypothesis revisited, *Int. Rev. Cytol.* 141 (1992) 233–357.
- [47] L. Margulis, Archaeal-eubacterial mergers in the origin of Eukarya: phylogenetic classification of life, *Proc. Natl. Acad. Sci. U. S. A.* 93 (3) (1996 Feb 6) 1071–1076.
- [48] L. Margulis, M.F. Dolan, R. Guerrero, The chimeric eukaryote: origin of the nucleus from the karyomastigont in amitochondriate protists, *Proc. Natl. Acad. Sci. U. S. A.* 97 (13) (2000 Jun 20) 6954–6959.
- [49] W. Martin, A briefly argued case that mitochondria and plastids are descendants of endosymbionts, but that the nuclear compartment is not, *Proc. R. Soc. B Biol. Sci.* 266 (1426) (1999 Jul 7) 1387.
- [50] D. Moreira, P. Lopez-Garcia, Symbiosis between methanogenic archaea and delta-proteobacteria as the origin of eukaryotes: the syntrophic hypothesis, *J. Mol. Evol.* 47 (5) (1998 Nov) 517–530.
- [51] R.A. Raff, H.R. Mahler, The non symbiotic origin of mitochondria, *Science* 177 (4049) (1972 Aug 18) 575–582.
- [52] T. Gabaldón, M.A. Huynen, Shaping the mitochondrial proteome, *Biochim. Biophys. Acta* 1659 (2–3) (2004 Dec 6) 212–220.
- [53] A. Bodyl, P. Mackiewicz, J.W. Stiller, Early steps in plastid evolution: current ideas and controversies, *BioEssays News Rev. Mol. Cell. Dev. Biol.* 31 (11) (2009 Nov) 1219–1232.
- [54] T. Cavalier-Smith, Membrane heredity and early chloroplast evolution, *Trends Plant Sci.* 5 (4) (2000 Apr) 174–182.
- [55] N. Duhita, H.A.T. Le, S. Satoshi, H. Kazuo, M. Daisuke, S. Takao, The origin of peroxisomes: the possibility of an actinobacterial symbiosis, *Gene* 450 (1–2) (2010 Jan 15) 18–24.