[Frontiers in Bioscience, Landmark, 22, 722-731, January 1, 2017]

The long story of mitochondrial DNA and respiratory complex I

Mauro Degli Esposti[1,2]

[1]*Centre for Genomic Sciences, UNAM, Cuernavaca, Morelos, Mexico,* [2]*Italian Institute of Technology, Genoa, Italy*

TABLE OF CONTENTS

1. ABSTRACT

This article examines the long story of the relationship between mitochondrial DNA (mtDNA) and respiratory complex I, NADH: Ubiquinone Oxidoreductase, from its beginning in the genome of the bacterial endosymbiont which then evolved into the mitochondria of our cells. The story begins with the evolution of ancient forms of bacterial complex I into the Nuo14 complex I that was present in the alpha proteobacterial ancestor of mitochondria. The story then becomes complicated in the diversity of eukaryotic organisms that are currently recognized. Therefore, it does not have a clear end, because currently available information shows different situations of metabolic adaptation and gene loss, indicating cases of de-evolution of the original protonmotive complex into a system that may fundamentally assist [FeFe]-hydrogenases in re-oxidising metabolically produced NADH under anaerobic conditions. The history of complex I is thus a never ending story of molecular and physiological evolution producing new perspectives for studying the enzyme complex that occupies the largest proportion of mitochondrial DNA.

2. INTRODUCTION

Respiratory complex I, NADH: Ubiquinone Oxidoreductase (Nuo), is the largest and most ancient proton pumping enzyme of the mitochondrial respiratory chain (1,2). Early biochemical studies with the enzyme complex purified from beef mitochondria defined three fractions enriched in the flavo-protein, iron-sulfur protein- and membrane protein-subunits (3). The membrane fraction was later found to comprise all seven products of the NADH Dehydrogenase (ND) genes encoded in the DNA of mammalian mitochondria (ND1-6 plus ND4L; 1-3). Subsequent studies revealed that the mitochondrial DNA (mtDNA) of plants, algae and protists additionally contains genes coding for hydrophilic redox subunits of the complex, from Nad7 (corresponding to beef 49kDa subunit; 3,4) to Nad11 (5,6,7). These hydrophilic subunits were found to be homologous to subunits of bacterial enzymes such as NiFe-hydrogenases (3,8,9), while the largest membrane subunits showed homologies to mrp (multiple resistance to pH) cation-proton antiporters (2,9,10,11). The modular architecture of complex I thus evolved by progressive addition of modules derived or related to different bacterial systems (1,2,11,12). The last stage of the evolution of this complex in bacteria produced an operon that contains 14 genes, in a sequence that recapitulates the addition of the NADH-reacting N-module to the ancestral Q-reacting module related to NiFe-hydrogenases and the membrane, proton pumping P module related to cation antiporters (Figure 1A and Table 1).

The mitochondria of our cells originated from a selected group of alpha proteobacteria that contained this Nuo14 operon together with the operons for succinate dehydrogenase (complex II), cytochrome *c* reductase (complex III o bc_1 complex) and cytochrome *c* oxidase (complex IV), which constitute the respiratory chain of biochemistry textbook (12,13). The most hydrophobic membrane subunits of these complexes have been retained in mtDNA. The genes for the majority of the hydrophilic and accessory subunits of the same complexes migrated into the nuclear DNA soon after the symbiogenic event that produced the progenitor of our cells, commonly defined as LECA - Last Eukaryotic Common Ancestor (7,12,14,15). Although only the cytochrome *b* subunit of complex III and the catalytic COX1 subunit of complex IV are present in all mtDNA

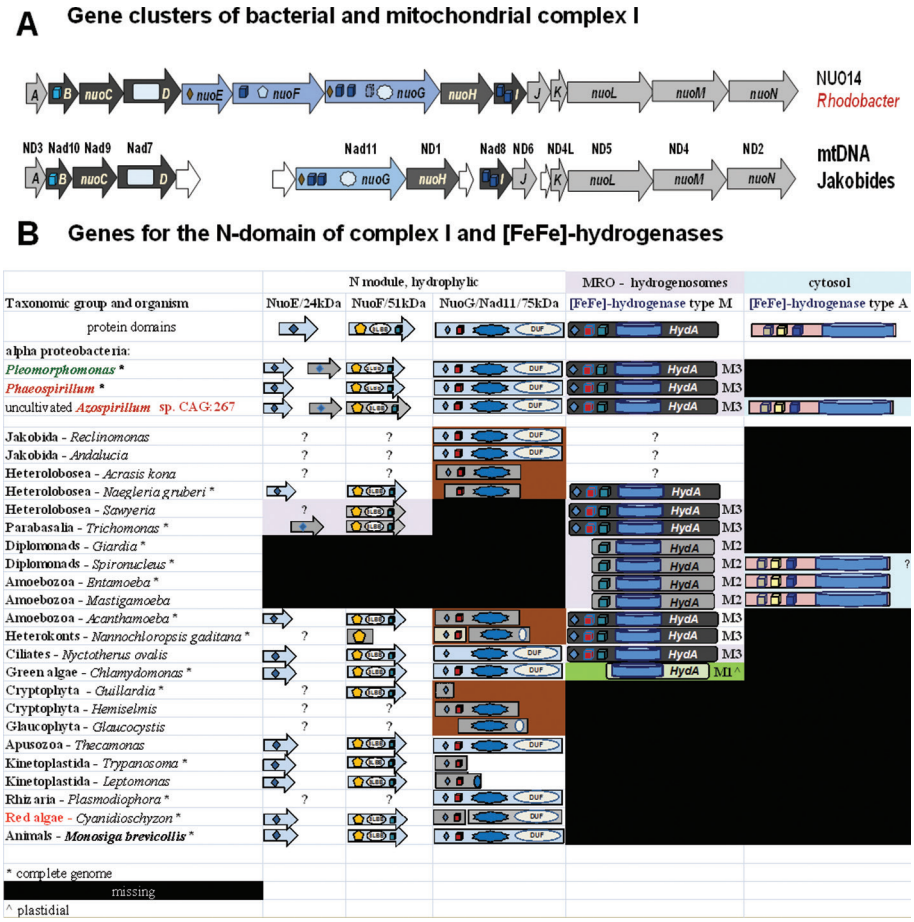

Figure 1. A. Gene sequence of a typical Nuo14 complex I from bacteria (20) compared with that of the mtDNA of the Excavata *Andalucia* (7). Proteins belonging to the Q-module derived from NiFe-hydrogenases are coloured in dark gray while those belonging to the membrane P-domain are in light gray (12). The three subunits forming the N-domain are in coloured in light blue. Empty arrows indicate non conserved genes that are not present in bacterial operons. B. N module of complex I and related [FeFe]-hydrogenases in aerobic and anaerobic eukaryotes. The recognized conserved domains (35) of the various proteins are rendered with the symbols previously used (12,19,30). The brown background in the column of the NuoG subunit indicates when it is encoded by mtDNA as Nad11. The question mark indicates that the gene is not found yet in currently available genomic information. The asterisk indicates organisms for which a complete genome is available now. The word 'missing' in white over black background indicates that the gene is missing in currently available genomes.

genomes sequenced so far (7,14,15), the genes for the subunits of complex I constitute a considerable proportion of the protein coding part of such genomes. This article will examine the variations in presence and composition of complex I subunits in the mtDNA of various protists in relation to the ancestral bacterial operon for the enzyme complex and draw functional conclusion on the evolution as well the de-evolution of this important bioenergetic system in mitochondria.

3. RESPIRATORY COMPLEX I AND MITOCHONDRIA

3.1. Ancient history of complex I

The giant enzyme complex that constitutes the respiratory complex I of mammalian mitochondria (1,2) has an ancient history, not just in evolutionary terms but also in functional transitions. The primordial form of the complex is found in strictly anaerobic bacteria that thrive in oceanic hydrothermal vents such as *Nautilia* (12), a deep branching epsilon proteobacterium (16). In these organisms, the operon of the enzyme lacks the NADH-reacting N-domain and may not even react with ubiquinone (Q), catalyzing instead electron transport between ferredoxins and NAD(P)H as in precursor forms of complex I such as *mbx*-like enzymes (12,17). Key changes in the NuoD and NuoB subunits, corresponding to Nad7 and Nad10 in the mtDNA of protists (Table 1), were responsible for the acquisition of Q reactivity in subsequent versions of complex I that are found in green sulfur bacteria such as *Chlorobium* (12). The assembly of the N-module to the *Chlorobium* type of Q-reacting complex I is likely to have occurred in anaerobic, sulfate oxidising proteobacteria of the delta class, because the genomes of these bacteria show the largest variation in complex I forms and gene cluster (12,17). In the genome

Table 1. mtDNA encoded subunits of complex I, Q and P modules

P module, membrane								Q module, hydrophylic			
Taxonomic group and organism	**NuoH/ ND1**	**NuoN/ ND2**	**NuoA/ ND3**	**NuoM/ ND4**	**NuoK/ ND4L**	**NuoL/ ND5**	**NuoJ/ ND6**	**NuoD/ Nad7**	**NuoI/ Nad8**	**NuoC/ Nad9**	**NuoB/ Nad10**
Jakobida - *Reclinomonas*	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Jakobida - *Andalucia*	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Tsukubamonas	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	U
Malawimonas	Y	Y	Y	Y	Y	Y	Y	U	U	Y	U
Heterolobosea - *Acrasis kona*	Y	Y	Y	Y	Y	Y	Y	Y	Y	U	U
Heterolobosea - *Naegleria gruberi*[1]	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	U
Kinetoplastida - *Trypanosoma*[1]	U	U	U	Y	U	Y	Y	S	Y	Y	N
Parabasalia - *Trichomonas*[1]	M	M	M	M	M	M	M	M	M	M	M
Diplomonads - *Giardia*[1]	M	M	M	M	M	M	M	M	M	M	M
Amoebozoa - *Entamoeba*[1]	M	M	M	M	M	M	M	M	M	M	M
Amoebozoa - *Mastigamoeba*	M	M	M	M	M	M	M	M	M	M	M
Cryptophyta - *Hemiselmis*	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Cryptophyta - *Rhodomonas*	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Heterokonts - *Nannochloropsis gaditana*[1]	Y	Y	Y	Y	Y	Y	Y	Y	U	Y	U
Heterokonts - *Heterosigma*[1]	Y	Y	Y	Y	Y	Y	Y	Y	U	Y	U
Ciliates - *Nyctotherus ovalis*	Y	Y	Y	Y	Y	Y[2]	U	Y	U	Y	U
Glaucophyta - *Glaucocystis*	Y	Y	Y	Y	Y	Y	Y	Y	U	Y	U
Rhizaria - *Bigelowiella*[1]	Y	Y	Y	Y	Y	Y	U	Y	U	Y	U
Rhizaria - *Spongospora*	Y	Y	Y	Y	Y	Y	Y	Y	U	Y	U
Amoebozoa - *Acanthamoeba*[1]	Y	Y	Y	Y	Y	Y	Y	Y	N	Y	N
Amoebozoa - *Verbamoeba*[1]	Y	Y	Y	Y	Y	Y	Y	Y	N	Y	N
Apusozoa - *Thecamonas*	Y	Y	Y	Y	Y	Y	Y	Y	N	N	N
Green algae - *Chlamydomonas*[1]	Y	Y	Y	Y	Y	Y	Y	Y	N	N	N
Red algae - *Cyanidioschyzon*[1]	Y	Y	Y	Y	Y	Y	Y	S	N	N	N
Animals - *Monosiga brevicollis*[1]	Y	Y	Y	Y	Y	Y	Y	S	N	N	N
Animals - *Homo sapiens*[1]	Y	Y	Y	Y	Y	Y	Y	S	N	N	N

The brown background indicates that the protein subunit is encoded by mtDNA. Yes: Y, U: Unknown, M: Missing, Nuclear: N, S: 49 kDa, [1]: Complete genome, [2]: Altered

of at least one delta proteobacterium, *Syntrophobacter* [accession Sfum_1954-Sfum_0202, cf (12)], the precursor form of the N-module deriving from NAD-dependent formate dehydrogenase is positioned upstream to the gene cluster for a *Chlorobium* type of complex I, thereby showing the addition of the last structural model of respiratory complex I in its genetic make. The final forms of bacterial complex I were thus produced, in two parallel gene clusters labeled Nuo13 and Nuo14 from the number of their constitutive subunits (12,18).

Nuo13 complex I is characterized by the fusion of the NuoC subunit, corresponding to Nad9 in the mtDNA of protists (Table 1), with the NuoD (Nad7) subunit. It also has an additional FeS cluster, cluster N7, in the NuoG subunit which corresponds to Nad11 in the mtDNA of protists (Figure 1). This form of the complex is characteristic of *Escherichia coli* and other gamma proteobacteria (8,17,18), but is present also in delta and epsilon proteobacteria, Nitrospirales and a group of alpha proteobacteria including *Rhodopseudomonas* and members of the Acetobacteraceae family (12,18). Some of the latter organisms also possess a Nuo14 complex I, in which there is no cluster N7 (except, perhaps, in the deep branching alpha proteobacterium, *Magnetococcus*) (12,18,19,20). The prototypic Nuo14

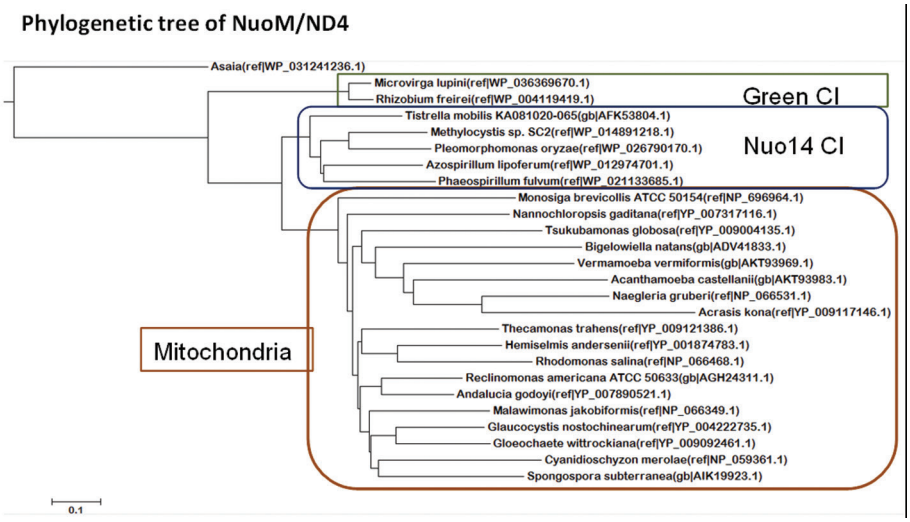

Figure 2. Phylogenetic tree of the NuoM/ND4 subunit of complex I. The tree represents the typical phylogenetic pattern of the membrane subunits of complex I from alpha proteobacteria and mitochondria from various eukaryotic lineages, cf. Table 1. The Neighbour-Joining (NJ) tree was obtained with DELTABLAST (39) search extended to 1000 sequences using NADH dehydrogenase subunit M (corresponding to mitochondrial ND4) of *Tistrella mobilis*, accession AFK53804, as previously repported (19). The outgroup is represented by the homologous NuoM of multispecies *Asaia* of the Acetobacteraceae family of alpha proteobacteria, which belong to Nuo13 complex I (12). Note the upstream position of two NuoM proteins representative of green complex I (green CI; 19). Note also that deep phylogenetic relationships between various eukaryotic groups are not well resolved using this proton pumping subunit of complex I, even if taxa of the same group generally cluster together, e.g. the Heterolobosea *Naegleria* and *Acrasis*.

complex I was discovered in the photosynthetic alpha proteobacterium, *Rhodobacter capsulatus* (20), but is also present in a variety of ancestral bacteria including *Thermus thermophilus* (12,18,19), of which the complete crystal structure has been obtained by Sazanov and coworkers (2,9). Presumably, Lateral Gene Transfer (LGT) from delta proteobacterial ancestors of Nuo14 complex I has spread this operon across various bacterial phyla first in anaerobic environments, from hydrothermal vents to the guts of termites (19,23), and subsequently also in (micro)aerobic but extreme environment such as those in which thermophilic *Thermus* thrive.

The crystal structure revealed the presence of cluster N7 in *Thermus* complex I (2), which therefore constitutes a relic of ancestral forms of the complex for it is present in the progenitor of the NuoG subunit, the Fds alpha subunit of NAD-dependent formate dehydrogenase (12,19,21). Recently, another form of bacterial complex I has been found in a restricted group of proteobacterial organisms, green complex I (19). It shows a different gene cluster that contains two additional genes, one for an antiporter membrane protein similar to NuoL (corresponding to the ND5 subunit coded by mtDNA) and a genetic regulator of the CBS family which is inserted between the genes for NuoG and NuoI (19). In phylogenetic trees, the subunits of green complex I occupy precursor or basal branches with respect to the subunits of alpha proteobacterial Nuo14 complexes (19), as shown here for the membrane subunit NuoM (corresponding to the ND4 subunit coded by mtDNA) (Figure 2). Conversely, NuoM and other subunits of the Nuo13 complex present in alpha proteobacteria such as *Asaia* occupy the most ancestral branch with respect to the homologous subunits of mitochondrial complex I (Figure 2).

To sum up, the ancient history of bacterial complex I culminated in different forms of the enzyme coded by three separate operons in alpha proteobacteria, from which the ancestors of mitochondria originated (13,22,23). These are: Nuo13, green complex I and Nuo14. However, only the Nuo14 operon was transmitted to mitochondria from the alpha proteobacterial ancestor, as clearly indicated by the presence and gene order of many complex I subunits in some protists (Figure 1A) (7,14,15,19,22).

3.2. Complex I just after eukaryogenesis and in extant protists

Eukaryotic cells evolved out of a single symbiogenic event that occurred when most sea environments were fundamentally anoxic (24). Consequently, the bacterium which then evolved into the mitochondrial organelle of eukaryotic cells must have provided not only the respiratory complexes and the catabolic systems characteristic of current mitochondria, but also enzyme systems typical of anaerobic metabolism, such as those involved in the extended glycolysis of anaerobic eukaryotes, for example *Entamoeba* (24,25). The signature enzyme of this anaerobic metabolism is [FeFe]-hydrogenase (25-27), a single subunit enzyme with a phylogenetic past that is entwined with that of respiratory complex I, especially in regard to its N-domain (Figure 1B).

Even before the elucidation of the 3D structure of the hydrophilic arm of respiratory complex I, a strong structural similarity had been noted between the N-terminal region the NuoG subunit of bacterial complex I and the equivalent region of the long form of [FeFe]-hydrogenase (type M3) (26). Similar to the related Fds-alpha subunit of formate dehydrogenase (21), this protein region contains a 2Fe2S ferredoxin followed by two 4Fe4S cluster (labeled cluster N4 and N5 in complex I nomenclature), the latter unusually ligated to one histidine and three cysteine residues (1,2). In some obligate anaerobic bacteria of the Clostridiales group, the gene for this [FeFe]-hydrogenase is associated with those of two proteins that are related to the NuoE and the NuoF subunits of complex I (21) forming an operon similar to that of NAD-dependent formate dehydrogenase (21). The eukaryotic orthologs of these proteins are the 24 and 51 kDa subunits of mitochondrial complex I (1,3) that are always coded by nuclear DNA in the eukaryotes described so far (7,14). The same proteins are found in the hydrogenosomes - a specialized form of Mitochondria Related Organelles (MRO; 24,25) adapted to anaerobiosis - of *Trichomonas* and the Heterolobosea *Sawyeria,* which lack complex I (25). These NADH-reacting subunits have been reported to originate from bacterial ancestors other than the alpha proteobacterial symbiont which transmitted the nuo operon of complex I to proto-eukaryotes (28). However, recently reported alpha proteobacteria from the human gut metagenome such as *Acetobacter* sp. CAG:977 (29) show divergent NuoE and NuoF subunits that cluster with the homologous proteins from *Trichomonas* and *Sawyeria*, as well as those fused within the gene for the [FeFe]-hydrogenase of the anaerobic ciliate *Nyctotherus* (30), rather than those of other alpha proteobacteria, as shown in Figure 3. Consequently, it is plausible that a single alpha proteobacterium related to *Acetobacter* sp. CAG:977 transmitted to LECA the anaerobic metabolism now found in eukaryotes such as *Trichomonas*, as well as the aerobic metabolism typical of animal mitochondria (30). It appears that this aerobic metabolism has been lost in extant organisms such as *Acetobacter* sp. CAG:977 as a consequence of the adaptation to the micro-oxic environment of our gut, and possibly also to symbiotic association with anaerobic eukaryotes.

The novel possibility that both the aerobic and the anaerobic metabolism of current eukaryotes derive from the same alpha proteobacterial ancestor of mitochondria (30) highlights the importance of respiratory complex I in the metabolic diversification that followed the evolution of LECA in the various taxonomic groups of protists and metazoans. This evolution was very rapid, producing a crown expansion of taxa that includes all the vast genomic variation that is recognized in the diverse groups of unicellular organisms classified so far (14,15,31-34). The present article will examine such a genomic variation following the changes in the distribution and structure of the subunits of complex I between the mitochondrial and nuclear genome (Table 1 and Figure 1B).

Although the root of the eukaryotic tree of life may be considered unresolved (31), converging evidence indicates that it may lie between two major taxonomic groups that do not contain plastids, namely Excavata and Unikonts - a term including the protists of the phylum Amoebozoa together with fungi and animals (31-33). Excavata group together free living and parasitic unicellular organisms generally containing flagella: predatory Jakobida (7), heterotrophic Heterolobosea (flagellate amoebas), parasitic Diplomonads (e.g. *Giardia*), parasitic Parabasalia (e.g. *Trichomonas*) and also Kinetoplastida, which include well known human pathogens such as *Trypanosoma* (31-33). The monophyletic nature of this collection of protists is uncertain (25,31), largely due to the fast evolving rate of obligate parasites such as *Giardia* and *Trypanosoma*, which have highly divergent proteins that normally skew phylogenetic trees of eukaryotes and their bacterial homologues (34). Nevertheless, Excavata display the largest genetic and molecular variation in Complex I subunits and also diverse types of [FeFe]-hydrogenases, as shown in Figure 1B. Although other taxonomic groups of eukaryotes contain [FeFe]-hydrogenases, including green and heterokont algae (25), only Amoebozoa show a variety of types for this enzyme matching that found among Excavata (Figure 1B). With regard to complex I, however, only Jakobida and the plastid-containing Cryptophyta have all the subunits ND1 to Nad10 encoded in their mtDNA (Table 1), thereby suggesting that the migration of some of these genes to nuclear DNA might have occurred in multiple separate events within different eukaryotic lineages.

The list of genes shown in Table 1, updating a previous study (14), suggests one interesting observation. Among bacterial Nuo subunits that have remained in the genome of mitochondrial DNA, NuoD (Nad7) and NuoC (Nad9) appear to be the most persistent, presumably reflecting their closeness in gene and protein structure in the bacterial enzyme, being fused together in ancestral Nuo13 complex I (see above section 3.1). Conversely, NuoB (Nad10) appears to be the least represented among the subunits encoded by mtDNA (Table 1). However, NuoG (Nad11) shows by far the largest variation in mtDNA encoding and structure of all the subunits of complex I inherited from bacteria (Figure 1B), as discussed in detail below.

3.3 Functional diversification and de-evolution of respiratory complex I

Bacterial NuoG originates from the Mo-containing largest subunit of NAD-dependent formate dehydrogenase (Fds alpha) (21,23). During its divergence into the homologous 75 kDa protein of mitochondrial complex I (1-3) it has lost two FeS clusters that were originally present in Fds-alpha and

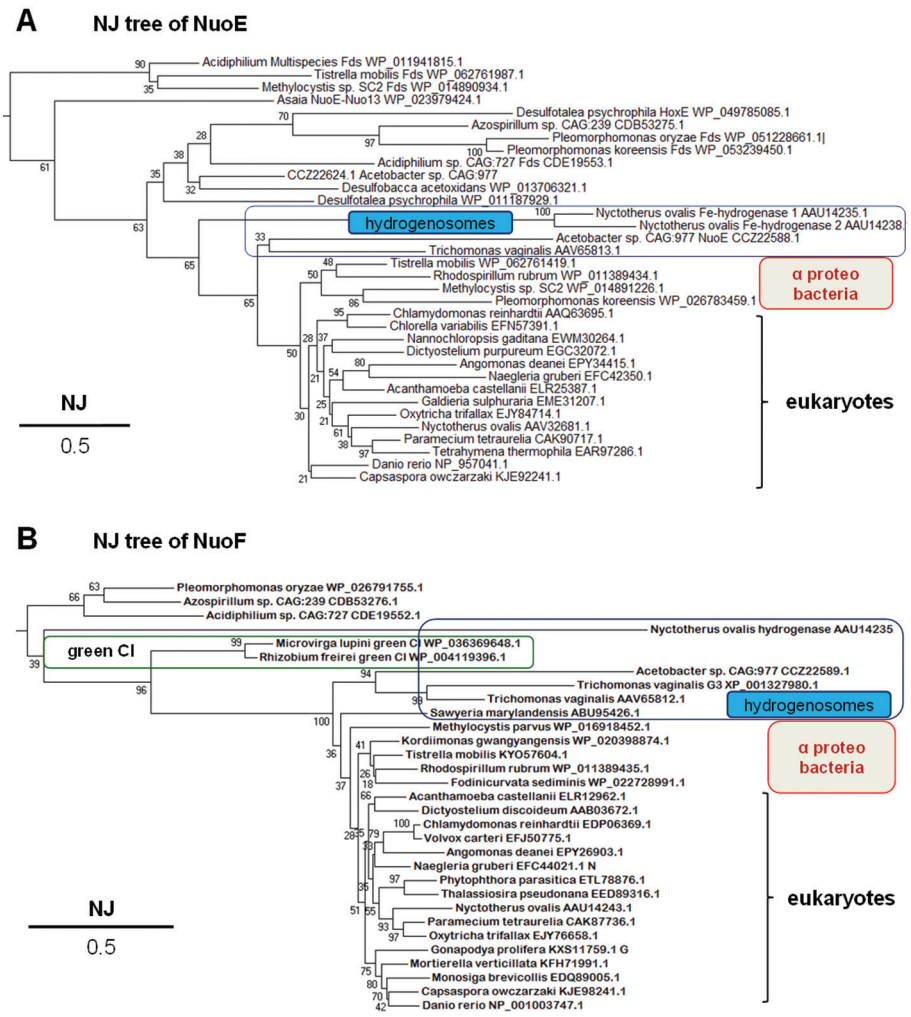

Figure 3. Phylogenetic trees for NuoE nd NuoF. Phylogenetic trees of complex I subunits of the NADH-reacting modules show clustering of hydrogenosomal proteins with *Acetobacter* sp. CAG:977. A. The NJ tree of the NuoE/24kDa subunit of complex I was obtained using a selection of proteins including those present in uncultured Rhodospirillales retrieved from wide DeltaBLAST searches. The protein sequences were first aligned using the CLUSTAW option of the MEGA 5.2 program (40) and subsequently manually refined taking into account the known 3D structure of the homologous proteins in *Thermus* complex I (2) as previously described (12). The manually refined alignment was then used for building a NJ tree with 200 bootstraps, the value of which are reported aside the nodes. The large box highlights the clustering of the NuoE subunit of *Acetobacter* sp. CAG:977 (29,30) with 24kDa subunit of hydrogenosomal origin. B. The NJ tree of the NuoF/51kDa subunit of complex I was obtained using a selection of proteins as in A. The box highlights the clustering of the NuoF subunit of *Acetobacter* sp. CAG:977 (29) with 51kDa subunits of hydrogenosomal origin. Homologous NuoF subunits of green complex I [green CI (19)] are enclosed in the smaller box with green rim. An outgroup protein from *Phaeospirillum* is cut off from the tree view.

subsequently retained in different types of bacterial complex I, as presented above. The loss of these clusters has streamlined the electron flow within the subunit, which connects the NuoF/51 kDa subunit containing the flavin cofactor that acts as first electron acceptor from the NADH substrate of complex I, to the twin FeS clusters contained in the NuoI subunit and ultimately to high potential cluster N2 contained in the NuoB/20 kDa subunit, which then functions as the final electron acceptor before electron transport to the Q substrate (1,2). All this electron transfer is catalysed by the FeS clusters that are present in 75 kDa subunit of mitochondrial complex I (1,2). However, the central and C-terminal part of the protein contain two domains that are shared with bacterial NuoG: the central, large Mo-binding domain [MopB_Res-Cmplx1_Nad11 according to the CDD classification (35), also labeled cd02773] and a Domain of Unknown Function (36), DUF1982, at the C-terminus. In the 3D structure of bacterial and mitochondrial complex I, these domains contribute to the stability of the overall hydrophilic arm of the complex but not to the electron transport function (2). Although a mitochondrial complex lacking these domains of the75 kDa subunit may be capable of electron transport from NADH to Q, it is unlikely to retain the overall stability of the native enzyme and therefore may be labile or partially functional at the physiology level of the cell.

Given the above considerations, it is remarkable that the Nad11 gene in the mitochondrial DNA of several protists does not code for the C-terminal DUF1982 domain of the 75 kDa subunit, while in other protists it does not have the N-terminal domain containing either cluster N4 or N5, or both (Figure 1B). Consequently, the presence of a Nad11 gene does not necessarily mean that the coded subunit is functional, and therefore that the whole complex I has retained its normal protonmotive function of transporting electrons from NADH to Q. This is particularly evident for *Trypanosoma* and other pathogenic Kinetoplastida (Figure 1B, Table 1 and data not shown), which may lack a functional protonmotive complex I in their mitochondria - even if complex I activity has been reported for a kinetoplastid parasite of plants (38). A similar conclusion can be drawn for Cryptophyta and Glaucophyta, ancient groups of algae which lack different structural elements of the native 75kDa subunit in their Nad11 gene (Figure 1B). Intriguingly, Heterolobosea such as *Naegleria* also have a truncated Nad11 gene that is unlikely to produce a functional 75kDa subunit, thereby suggesting that their complex I is either partially inactive, or fulfills other roles in mitochondria. However, *Naegleria* also has the gene for a fully functional [FeFe]-hydrogenase related to NuoG (Figure 1B), suggesting that this protein may be functionally linked to the incomplete complex I, substituting its NADH re-oxidising function with a diversion of electron from membrane-bound Q to the production of soluble H_2. Such a function is actually accomplished in the hydrogenosomes of the related Heterolobosea *Sawyeria* (25,37), which contains the reduced version of the N-domain of complex constituted by the 24 and 51 kDa subunits I as in the hydrogenosomes of *Trichomonas* (25,28). Indeed, both *Sawyeria* and *Trichomonas* lack other genes for complex I subunits (Figure 1B and Table 1) and therefore do not have the enzyme altogether (25).

The aerobic Amoebozoa, *Acanthamoeba castellani*, seems to be similar to *Naegleria* in this regard, since it has a [FeFe]-hydrogenase related to NuoG which may functionally compensate for its incomplete 75kDa subunit lacking the C-terminal DUF domain (Figure 1B). Indeed, both *Naegleria* and *Acanthamoeba* have been reported to adapt to micro-oxic conditions by producing hydrogen via this hydrogenase (25). Therefore, members of the Excavata group such as *Naegleria* and of the Amoebozoa such as *Acanthamoeba* have similarly retained the metabolic versatility of switching from aerobic mitochondrial metabolism. Under aerobic conditions, complex I may have only limited functional capacity, whereas under anaerobic conditions, the same mitochondrial complex I may assist electron diversion from NADH to hydrogen as in classical hydrogenosomes. This scenario may well reflect the ancestral metabolic versatility that has been transmitted to LECA by the alpha proteobacterial progenitor of mitochondria (30), which points out once more to the ancestry of all eukaryotes being rooted between Excavata and Amoebozoa (31,33,34).

4. CONCLUSIONS AND PERSPECTIVES

In this article I have summarized the long story of the relationship between mitochondrial DNA and respiratory complex I, the subunit of which contribute a large portion of the protein-coding part of the mtDNA in mammals, as well as the majority of eukaryotes (14). The story begun with the bacterial Nuo14 complex I that was present in the alphaproteobacterial ancestor of mitochondria. This complex is likely to have also contributed elements of anaerobic metabolism that are present in extant eukaryotes and connected, both structurally and functionally, to respiratory complex I. The analysis of the distribution and structure of complex I subunits undertaken here (cf. Figure 2) generates new intriguing possibilities regarding the true physiological function of the enzyme in several taxonomic groups of protists and algae that need experimental verification in the future. These possibilities remain tentative at the moment, due to the incomplete genomic information that is currently available for key subunits of complex I in major taxonomic groups of eukaryotes (Figure 1B). Filling in the missing genomic information will provide a clearer picture of the structure and function of complex I and related metabolic pathways pivoting on NADH in the variety of eukaryotic lineages that are presently known. This information will also elucidate whether the incomplete structures of the crucial 75kDa subunit of the enzyme complex presented here (Figure 1B) reflect real cases of de-evolution of complex I into inactive or redundant forms of the enzyme, or may instead represent adaptations that recapitulate the ancestral metabolic versatility of the very first mitochondrion and its bacterial ancestor.

5. ACKNOWLEDGEMENTS

This work was sponsored by intramural funds at IIT and by CONACyT grant No. 263876 for the sabbatical research period of MDE in Mexico.

6. REFERENCES

1. U Brandt: Energy converting NADH: quinone oxidoreductase (complex I). *Annu Rev Biochem* 75, 69–92 (2006) DOI: 10.1146/annurev.biochem.75.103004.142539 PMid:16756485

2. LA Sazanov: A giant molecular proton pump: structure and mechanism of respiratory complex I. *Nat Rev Mol Cell Biol* 16, 375-388 (2015) DOI: 10.1038/nrm3997

PMid:25991374

3. IM Fearnley, JE Walker: Conservation of sequences of subunits of mitochondrial complex I and their relationships with other proteins. *Biochim Biophys Acta* 1140, 105–134 (1991)
DOI: 10.1016/0005-2728(92)90001-I

4. L Bonen L, K Williams K, S Bird S, C Wood C: The NADH dehydrogenase subunit 7 gene is interrupted by four group II introns in the wheat mitochondrial genome. *Mol Gen Genet* 244,81-89 (1994)
DOI: 10.1007/BF00280190

5. EV Armbrust, JA Berges, C Bowler, BR Green, D Martinez, NH Putnam, S Zhou, AE Allen, KE Apt, M Bechner, MA Brzezinski, BK Chaal, A Chiovitti, AK Davis, MS Demarest, JC Detter, T Glavina, D Goodstein, MZ Hadi, U Hellsten, M Hildebrand, BD Jenkins, J Jurka, VV Kapitonov, N Kröger, Lau WW, TW Lane, L FW arimer, JC Lippmeier, S Lucas, M Medina, A Montsant, M Obornik MS, Parker, B Palenik, GJ Pazour, PM Richardson, TA Rynearson, MA Saito, DC Schwartz, K Thamatrakoln, K Valentin, A Vardi, FP Wilkerson, DS Rokhsar. The genome of the diatom Thalassiosira pseudonana: ecology, evolution, and metabolism. *Science* 306, 79-86 (2004)
DOI: 10.1126/science.1101156
PMid:15459382

6. BA Curtis, G Tanifuji, F Burki, A Gruber, M Irimia, S Maruyama, MC Arias, SG Ball, GH Gile, Y Hirakawa, JF Hopkins, A Kuo, SA Rensing, J Schmutz, A Symeonidi, M Elias, RJ Eveleigh, EK Herman, MJ Klute, T Nakayama, M Oborník, A Reyes-Prieto, EV Armbrust, SJ Aves, RG Beiko, P Coutinho, JB Dacks, DG Durnford, NM Fast, BR Green, CJ Grisdale, F Hempel, B Henrissat, MP Höppner, K Ishida, E Kim, L Kořený, PG Kroth, Y Liu, SB Malik, UG Maier, Mc D Rose, M T ock, JA Neilson, NT Onodera, Poole AM, EJ Pritham, TA Richards, G Rocap, SW Roy, C Sarai, S Schaack, S Shirato, CH Slamovits, DF Spencer, S Suzuki, AZ Worden, S Zauner, K Barry, C Bell, AK Bharti, JA Crow, J Grimwood, R Kramer, E Lindquist, S Lucas, A Salamov, GI McFadden, CE Lane, PJ Keeling, MW Gray, IV Grigoriev, JM Archibald. Algal genomes reveal evolutionary mosaicism and the fate of nucleomorphs. *Nature* 492, 59-65 (2012)
DOI: 10.1038/nature11681
PMid:23201678

7. G Burger, MW Gray, L Forget, BF Lang: Strikingly bacteria-like and gene-rich mitochondrial genomes throughout jakobid protists. *Genome Biol Evol* 5, 418-38 (2012)
DOI: 10.1093/gbe/evt008
PMid:23335123 PMCid: PMC3590771

8. T Friedrich, D Scheide: The respiratory complex I of bacteria, archaea and eukarya and its module common with membrane-bound multisubunit hydrogenases. *FEBS Lett* 479, 1–5 (2000)
DOI: 10.1016/S0014-5793(00)01867-6

9. RG Efremov, LA Sazanov: The coupling mechanism of respiratory complex I — a structural and evolutionary perspective. *Biochim Biophys Acta* 1817, 1785–1795 (2012)
DOI: 10.1016/j.bbabio.2012.02.015
PMid:22386882

10. C Mathiesen, C Hagerhall: Transmembrane topology of the NuoL, M and N subunits of NADH: quinone oxidoreductase and their homologues among membrane-bound hydrogenases and bona fide antiporters. *Biochim Biophys Acta* 1556, 121–132 (2002)
DOI: 10.1016/S0005-2728(02)00343-2

11. VK Moparthi, C Hagerhall: The evolution of respiratory chain complex I from a smaller last common ancestor consisting of 11 protein subunits. *J Mol Evol* 72, 484-497 (2011)
DOI: 10.1007/s00239-011-9447-2
PMid:21597881 PMCid: PMC3144371

12. M Degli Esposti M: Genome Analysis of Structure-Function Relationships in Respiratory Complex I, an Ancient Bioenergetic Enzyme. *Genome Biol Evol* 8, 126-47 (2015)
DOI: 10.1093/gbe/evv239
PMid:26615219 PMCid: PMC4758237

13. M Degli Esposti, B Chouaia, F Comandatore, E Crotti, D Sassera, PM Lievens, D Daffonchio, C Bandi: Evolution of mitochondria reconstructed from the energy metabolism of living bacteria. *PLoS One* 9, e96566 (2014).
DOI: 10.1371/journal.pone.0096566

14. S Kannan, IB Rogozin, EV Koonin: MitoCOGs: clusters of orthologous genes from mitochondria and implications for the evolution of eukaryotes. *BMC Evol Biol* 14, 237 (2014)

15. RM de Graaf, G Ricard, TA van Alen, I Duarte, BE Dutilh, C Burgtorf, JW Kuiper, GW van der Staay, AG Tielens, MA Huynen, JH Hackstein: The organellar genome and metabolic potential of the hydrogen-producing mitochondrion of Nyctotherus ovalis. *Mol Biol Evol* 28,2379-91 (2011)
DOI: 10.1093/molbev/msr059
PMid:21378103 PMCid: PMC3144386

16. BJ Campbell, JL Smith, TE Hanson, MG Klotz, LY Stein, CK Lee, D Wu, JM Robinson, HM Khouri, JA Eisen, SC Cary: Adaptations to submarine hydrothermal environments exemplified by the genome of Nautilia profundicola. *PLoS Genet* 5, e1000362 (2009)

17. BC Marreiros, AP Batista, AM Duarte, MM Pereira: A missing link between complex I and group 4 membrane-bound [NiFe] hydrogenases. *Biochim Biophys Acta* 1827, 198-209 (2013)
DOI: 10.1016/j.bbabio.2012.09.012
PMid:23000657

18. MA Spero FO, Aylward, CR Currie, TJ Donohue: Phylogenomic analysis and predicted physiological role of the proton-translocating NADH: quinone oxidoreductase (complex I) across bacteria. *Mbio* 6, e00389-15 (2015)
DOI: 10.1128/mBio.00389-15
PMid:25873378 PMCid: PMC4453560

19. M Degli Esposti M, E Martinez Romero: A survey of the energy metabolism of nodulating symbionts reveals a new form of respiratory complex I. *FEMS Microbiol Ecol* 92(6) (2016)

20. A Dupuis, M Chevallet, E Darrouzet, H Duborjal, J Lunardi, JP Issartel: The complex I from Rhodobacter capsulatus. *Biochim Biophys Acta* 1364, 147–165 (1998)
DOI: 10.1016/S0005-2728(98)00025-5

21. JI Oh, B Bowien: Structural analysis of the fds operon encoding the NAD+-linked formate dehydrogenase of Ralstonia eutropha. *J Biol Chem* 273, 26349-26360 (1998)
DOI: 10.1074/jbc.273.41.26349
PMid:9756865

22. MW Gray: Mosaic nature of the mitochondrial proteome: Implications for the origin and evolution of mitochondria. *Proc Natl Acad Sci U S A*, 12:10133-8. (2015)

23. M Degli Esposti: Bioenergetic evolution in proteobacteria and mitochondria. Genome Biol Evol 6, 3238-51 (2016)
DOI: 10.1093/gbe/evu257
PMid:25432941

24. M Müller, M Mentel, JJ van Hellemond, K Henze, C Woehle, SB Gould, RY Yu, M van der Giezen, AG Tielens, WF Martin: Biochemistry and evolution of anaerobic energy metabolism in eukaryotes. *Microbiol Mol Biol Rev* 6:444-495 (2012)

25. CW Stairs, MM Leger, AJ Roger: Diversity and origins of anaerobic metabolism in mitochondria and related organelles. *Philos Trans R Soc Lond B Biol Sci* 370,20140326. (2015)

26. J Meyer: [FeFe] hydrogenases and their evolution: a genomic perspective. *Cell Mol Life Sci* 64:1063-84 (2007)

27. A Graentzdoerffer, D Rauh, A Pich, JR Andreesen: Molecular and biochemical characterization of two tungsten- and selenium-containing formate dehydrogenases from Eubacterium acidaminophilum that are associated with components of an iron-only hydrogenase. *Arch Microbiol* 179,116-30 (2003)

28. SD Dyall, W Yan, MG Delgadillo-Correa, A Lunceford, JA Loo, CF Clarke, PJ Johnson: Non-mitochondrial omplex I proteins in a hydrogenosomal oxidoreductase complex. *Nature* 431, 1103-7 (2003)
DOI: 10.1038/nature02990
PMid:15510149

29. HB Nielsen, M Almeida, AS Juncker, S Rasmussen, J Li, S Sunagawa, DR Plichta, L Gautier, AG Pedersen, E Le Chatelier, E Pelletier, I Bonde, T Nielsen, C Manichanh, M Arumugam, JM Batto, MB Quintanilha Dos Santos, N Blom, N Borruel, KS Burgdorf, F Boumezbeur, F Casellas, J Doré, P Dworzynski, F Guarner, T Hansen, F Hildebrand, RS Kaas, S Kennedy, K Kristiansen, JR Kultima, P Léonard, F Levenez, O Lund, B Moumen, D Le Paslier, N Pons, O Pedersen, E Prifti, J Qin, J Raes, S Sørensen, J Tap, S Tims, DW Ussery, T Yamada; MetaHIT Consortium, P Renault, T Sicheritz-Ponten, P Bork, J Wang, S Brunak, SD Ehrlich; MetaHIT Consortium: Identification and assembly of genomes and genetic elements in complex metagenomic samples without using reference genomes.

Nat Biotechnol 32, 822-8 (2014)
DOI: 10.1038/nbt.2939
PMid:24997787

30. Degli Esposti M, Cortez D, Lozano L, Rasmussen S, Nielsen HB, Esperanza Martinez Romero E. Alpha proteobacterial ancestry of the [Fe-Fe]-hydrogenases in anaerobic eukaryotes. *Biol Direct* 11:34 (2016)
31. F Burki: The eukaryotic tree of life from a global phylogenomic perspective. *Cold Spring Harb Perspect Biol* 6, a016147 (2014)
32. D He, O Fiz-Palacios, CJ Fu, J Fehling, CC Tsai, SL Baldauf: An alternative root for the eukaryote tree of life. *Curr Biol* 24, 465-70 (2014).
DOI: 10.1016/j.cub.2014.01.036
PMid:24508168
33. R Derelle, G Torruella, V Klimeš, H Brinkmann, E Kim, Č Vlček, BF Lang, M Eliáš: Bacterial proteins pinpoint a single eukaryotic root. *Proc Natl Acad Sci U S A* 112, E693-9 (2015)
DOI: 10.1073/pnas.1420657112
PMid:25646484 PMCid: PMC4343179
34. F Burki, M Kaplan, DV Tikhonenkov, V Zlatogursky, BQ Minh, LV Radaykina, A Smirnov, AP Mylnikov, PJ Keeling: Untangling the early diversification of eukaryotes: a phylogenomic study of the evolutionary origins of Centrohelida, Haptophyta and Cryptista. *Proc Biol Sci* 283(1823) (2016)
35. A Marchler-Bauer, MK Derbyshire, NR Gonzales, S Lu, F Chitsaz, LY Geer, RC Geer, J He, M Gwadz, DI Hurwitz, CJ Lanczycki, F Lu, GH Marchler, JS Song, N Thanki, Z Wang, RA Yamashita, D Zhang, C Zheng, SH Bryant: CDD: NCBI's conserved domain database. *Nucleic Acids Res* 43(Database issue):D222-6 (2015)
36. NF Goodacre, DL Gerloff, P Uetz: Protein domains of unknown function are essential in bacteria. *Mbio* 5, e00744-13 (2013)
37. MJ Barberà, I Ruiz-Trillo, JY Tufts, A Bery, JD Silberman, AJ Roger: Sawyeria marylandensis (Heterolobosea) has a hydrogenosome with novel metabolic properties. *Eukaryot Cell* 9, 1913-24 (2010)
DOI: 10.1128/EC.00122-10
PMid:21037180 PMCid: PMC3008280
38. D González-Halphen, DA Maslov: NADH-ubiquinone oxidoreductase activity in the kinetoplasts of the plant trypanosomatid Phytomonas serpens. *Parasitol Res* 92, 341-6 (2004)
DOI: 10.1007/s00436-003-1058-4
PMid:14727190
39. GM Boratyn, AA Schäffer, R Agarwala, SF Altschul, DJ Lipman, TL Madden: Domain enhanced lookup time accelerated BLAST. *Biol Direct* 7, 12 (2012)
40. K Tamura, D Peterson, N Peterson, G Stecher, M Nei, S Kumar: MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol* 28, 2731-9 (2011)
DOI: 10.1093/molbev/msr121
PMid:21546353 PMCid: PMC3203626

Abbreviations: LECA, Last Eukaryotic Common Ancestor; LGT, Lateral Gene Transfer; NJ, Neighbor-Joining; Q, ubiquinone.

Key Words: Respiratory Complex I, Bacterial Evolution, Mitchondrial Evolution, Anaereobic Metabolism, Review

Send correspondence to: Mauro Degli Esposti, Centre for Genomic Sciences, UNAM, Cuernavaca, Morelos, Mexico and Italian Institute of Technology, Genoa, Italy, Tel: 0039 010 71781987, E-mail: mauro.degliesposti@iit.it