

Minireview

## What makes a mitochondrion?

Joshua L Heazlewood, A Harvey Millar, David A Day and James Whelan

Address: Plant Molecular Biology Group, School of Biomedical and Chemical Sciences, University of Western Australia, 35 Stirling Highway, Crawley 6009, Western Australia, Australia.

Correspondence: James Whelan. E-mail: [seamus@cyllene.uwa.edu.au](mailto:seamus@cyllene.uwa.edu.au)

Published: 28 May 2003

*Genome Biology* 2003, 4:218

The electronic version of this article is the complete one and can be found online at <http://genomebiology.com/2003/4/6/218>

© 2003 BioMed Central Ltd

### Abstract

Experimental analyses of the proteins found in the mitochondria of yeast, humans and *Arabidopsis* have confirmed some expectations but given some surprises and some insights into the evolutionary origins of mitochondrial proteins.

With the completion of the genome sequences of yeast, human and *Arabidopsis*, which contain approximately 6,000, 35,000 and 28,000 genes, respectively [1-3], the world's attention is now shifting to elucidation of gene function, and major proteomic studies are currently under way on a variety of organisms [4-6]. As a step towards assembling a list of the total complement of proteins in any one cell type (its proteome), proteomic studies of subcellular compartments and organelles have become a major focus, because smaller and more manageable subsets of proteins are involved. Given that compartmentation is a hallmark of the eukaryotic cell, and because the functions of organelles are biochemically well defined, such studies have an immediate functional impact, in contrast to the relatively limited insights that can be gained from the complete, unstructured cell proteome.

Mitochondria are attractive targets for subcellular proteomics because they play vital roles in energy production, anabolic and catabolic metabolism and in programmed cell death pathways, they can be purified readily from model organisms, and defects in mitochondrial proteins can have dramatic effects on the functions of cells and organs. Defining mitochondrial proteomes in a number of model organisms across the divisions of eukaryotes facilitates cross-species comparisons, thus greatly aiding validation of conclusions from each species and providing insights into both function and evolution [5].

The recent identification of 615 proteins from the mitochondrial proteome of the human heart [7] represents the first comprehensive analysis of a mitochondrial proteome and the highest number of proteins identified to date from any subcellular compartment. This is likely to change soon, as concerted efforts towards defining other subcellular proteomes are currently in progress [6,8]. We now have glimpses of the mitochondrial proteomes from the yeast *Saccharomyces cerevisiae* and *Arabidopsis*, as well as humans (Table 1), although these are far from complete. Various approaches have predicted that approximately 10% of the coding capacity of the nuclear genome is devoted to proteins destined for the mitochondrion [9-11]. For yeast, predictions of the total number of proteins in a mitochondrion, made using a combination of sequence homology and gene tagging or knockouts, vary between 423 and 630 proteins, which is close to the number predicted by a variety of bioinformatic analyses of protein targeting [9-11]. Direct protein sequencing using mass spectrometry has so far yielded only 179 mitochondrial proteins, however, and gene-tagging and knockout analysis have given 332 and 466 proteins, respectively [12,13]. Thus, even in yeast, the experimentally confirmed proteome is less than 50% complete, according to current predictions. In plants, the experimentally determined set so far contains only 135 mitochondrial proteins for *Arabidopsis* [14,15] and 136 for rice [16]; these numbers are significantly lower than the 10% of the nuclear genome that is predicted by bioinformatic approaches to encode mitochondrial proteins [3,9].

**Table 1****Predicted and experimentally determined numbers of proteins present in mitochondria**

	Yeast	Reference	Human	Reference	<i>Arabidopsis</i>	Reference
<b>Predictions*</b>						
Total	423	[29]	734	[17]	800	[20]
	584	[17]	1,500	[30]		
	630	[11]				
Prediction percentage <sup>†</sup>	10-13%	[10,11]	10%	[9]	10%	[9]
Extrapolated <sup>‡</sup>	617-802		3,500		2,800	
<b>Experimental</b>						
Total	179	[13]	615	[7]	135	[14,15,18,20,31]
	388	[32]				
	466	[12]				
Experimental percentage <sup>§</sup>	2.9%		1.76%		0.48%	
	6.3%					
	7.6%					

\*Predictions of the number of mitochondrial proteins are from sequence homology, targeting sequences, phylogenetic profiling or extrapolation from a set of experimental values. <sup>†</sup>Prediction percentage: the percentage of genes in the genome predicted to encode mitochondrial proteins from targeting analyses or phylogenetic profiles; <sup>‡</sup>Extrapolated: the total number of mitochondrial proteins predicted from the percentage value given and the genome size of each organism. <sup>§</sup>Experimental percentage: the percentage of the predicted proteome found in experimentally determined mitochondrial proteomes.

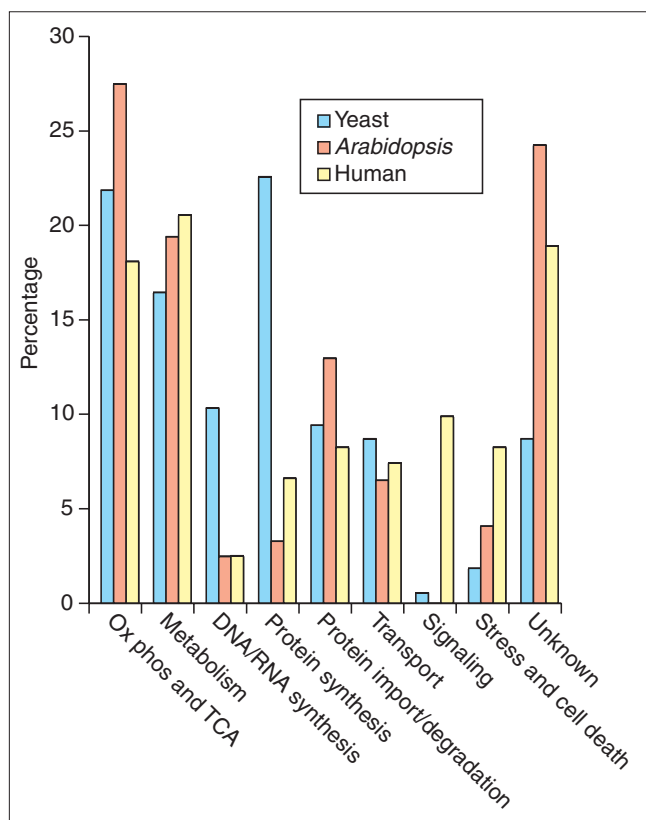
Even the 615 proteins directly identified in human mitochondria represent only about 25-35% of the proteins predicted to be mitochondrial by targeting analyses and by extrapolations from yeast studies [8,9]. In reality, the true number of mitochondrial proteins will probably lie somewhere between the current experimentally determined numbers and the predictions.

Sorting the identified sets of proteins (either predicted or known) by their functions reveals both expected and unexpected outcomes (Figure 1). Such comparisons vary slightly depending on the lists used, but those shown here are based on the functional analyses reported for *Arabidopsis* [14,15], human [7] and yeast [17]. The yeast protein set is derived from both genetic and mass-spectrometric data, whereas the human and *Arabidopsis* sets are derived only from mass spectrometry; this means that more low-abundance DNA-, RNA- and protein-synthesis components have been identified in yeast than in the other two species.

As expected, the predominant mitochondrial proteins found are oxidative-phosphorylation complexes, enzymes of the tricarboxylic acid cycle, components of the protein-import and protein-synthesis machinery, and transport proteins; these represent one third to one half of the identified sets in each species. The large number of proteins of unknown function (10-20%) and the large number of enzymes of the carbohydrate, amino-acid and lipid metabolism pathways have come as more of a surprise, however. In particular,

the presence of glycolytic enzymes in purified mitochondrial preparations, and the diverse kinds of predicted signaling components such as kinases and receptors, were largely unexpected, as their presence in mitochondria has not been documented in earlier studies. These findings need further substantiation, and this has become an area of active research, as has the search for protein-protein associations within the proteome [7,8,18-20]. The absence of some proteins is also perplexing. For example, despite the presence of many genes from the mitochondrial carrier superfamily in all of the genomes so far examined, only a handful of carrier proteins have been experimentally identified in mitochondria to date [7,18].

Mitochondrial proteomes also need to be defined in terms of their evolutionary origins. Mitochondria almost certainly evolved from an  $\alpha$ -proteobacterium that was engulfed by an early eukaryotic cell and entered into symbiosis with it. Surprisingly, conservative estimates indicate that, in yeast, only 25-50% of mitochondrial proteins can be identified as most closely related to  $\alpha$ -proteobacterial proteins [21,22]. This suggests that approaches to defining subcellular proteomes that rely on homology to prokaryotic 'ancestors' are useful but have limitations. Divergence of the mitochondrial proteomes between different major eukaryotic lineages may mean that, even in identical pathways, components in one organism may have different phylogenetic origins from the equivalent components in another [21]. A glimpse of this is seen with the mitochondrial ribosome of *Arabidopsis*, which has proteins



**Figure 1**

Functional classification of the proteins from the experimentally determined proteomes of yeast, *Arabidopsis* and human. (Ox phos, oxidative phosphorylation; TCA, tricarboxylic acid cycle).

from three distinct genetic origins: the mitochondrion, the plastid and the nucleus of the host eukaryotic cell [23].

It is evident that mitochondrial proteomes have undergone expansion in function during evolution, in addition to the loss of bacterial metabolic pathways such as glycolysis [21]. The evolutionary expansion of mitochondrial proteomes means that proteins of eukaryotic origin are also represented in the mitochondrial proteome, complicating comparisons with  $\alpha$ -proteobacterial ancestors [24]. In plants the situation is further complicated by proteins of cyanobacterial origin, presumably gained from chloroplasts via gene transfer from the plastid to the nucleus and subsequent duplication and re-targeting to mitochondria [23]. It has been observed that proteins derived from  $\alpha$ -proteobacteria that are found in mitochondria but encoded in the nucleus appear to be preferentially synthesized on ribosomes attached to the mitochondria [25]; this may provide an experimental avenue for investigating the different genetic origins of mitochondrial proteins.

From an evolutionary point of view, it is tempting to estimate the numbers of mitochondrial proteins by comparison with

modern-day obligate intracellular parasites, such as *Rickettsia prowazekii*, which contains 834 proteins [26]. Many common functions found in mitochondria, such as amino-acid biosynthetic pathways, are absent from these parasites, however. Obligate intracellular parasites provide examples of genome reduction, and the mitochondrial ancestor almost certainly had a larger genome and protein-coding capability than *Rickettsia*.

Defining the complete mitochondrial proteome will require a variety of experimental approaches, including the direct proteomic-identification and protein-tagging strategies that are presently underway [6]. Defining a static mitochondrial proteome will certainly be an achievement, but this is only the beginning. Determining how the proteome changes under certain conditions, such as during oxidative stress [27,28], between tissues and through development, will use this basic set of proteins as a platform. Identifying new functions and interactions of proteins, and of signal-transduction pathways, will require knockouts, overexpression experiments and analysis of the phosphorylated components of the proteome [5]. Finally, comparative mitochondrial proteomics between organisms will give insights into how proteins have diverged in function through evolution and may well help answer the still vexing question of the ancestral origins of the eukaryotic cell.

## References

- Goffeau A, Aert R, Agostini-Carbone ML, Aigle AM, Alberghina L, Albermann K, Albers M, Aldea M, Alexandraki D, Aljinoni G, et al.: **The yeast genome directory.** *Nature* 1997, **Suppl 287**:1-105.
- Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, Devon K, Dewar K, Doyle M, FitzHugh W, et al.: **Initial sequencing and analysis of the human genome.** *Nature* 2001, **409**:860-921.
- Arabidopsis Genome Initiative: **Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*.** *Nature* 2000, **408**:796-815.
- Koller A, Washburn MP, Lange BM, Andon NL, Deciu C, Haynes PA, Hays L, Schieltz D, Ulaszek R, Wei J, et al.: **Proteomic survey of metabolic pathways in rice.** *Proc Natl Acad Sci USA* 2002, **99**:11969-11974.
- Tyers M, Mann M: **From genomics to proteomics.** *Nature* 2003, **422**:193-197.
- Aebersold R, Mann M: **Mass spectrometry-based proteomics.** *Nature* 2003, **422**:198-207.
- Taylor SW, Fahy E, Zhang B, Glenn GM, Warnock DE, Wiley S, Murphy AN, Gaucher SP, Capaldi RA, Gibson BW, et al.: **Characterization of the human heart mitochondrial proteome.** *Nat Biotechnol* 2003, **21**:281-286.
- Taylor SW, Fahy E, Ghosh SS: **Global organellar proteomics.** *Trends Biotechnol* 2003, **21**:82-88.
- Emanuelsson O, Nielsen H, Brunak S, von Heijne G: **Predicting subcellular localization of proteins based on their N-terminal amino acid sequence.** *J Mol Biol* 2000, **300**:1005-1016.
- Kumar A, Agarwal S, Heyman JA, Matson S, Heidtman M, Piccirillo S, Umansky L, Drawid A, Jansen R, Liu Y, et al.: **Subcellular localization of the yeast proteome.** *Genes Dev* 2002, **16**:707-719.
- Marcotte EM, Xenarios I, van der Bliek AM, Eisenberg D: **Localizing proteins in the cell from their phylogenetic profiles.** *Proc Natl Acad Sci USA* 2000, **97**:12115-12120.
- Steinmetz LM, Scharfe C, Deutschbauer AM, Mokranjac D, Herman ZS, Jones T, Chu AM, Giaever G, Prokisch H, Oefner PJ, et al.: **Systematic screen for human disease genes in yeast.** *Nature Genet* 2002, **31**:400-404.
- Pflieger D, Le Caer JP, Lemaire C, Bernard BA, Dujardin G, Rossier J: **Systematic identification of mitochondrial proteins by LC-MS/MS.** *Anal Chem* 2002, **74**:2400-2406.

14. Krufft V, Eubel H, Jansch L, Werhahn W, Braun HP: **Proteomic approach to identify novel mitochondrial proteins in *Arabidopsis***. *Plant Physiol* 2001, **127**:1694-1710.
15. Millar AH, Sweetlove LJ, Giege P, Leaver CJ: **Analysis of the *Arabidopsis* mitochondrial proteome**. *Plant Physiol* 2001, **127**:1711-1727.
16. Heazlewood JL, Howell KA, Whelan J, Millar AH: **Towards an analysis of the rice mitochondrial proteome**. *Plant Physiol* 2003, **132**:230-242.
17. Schon EA: **Gene products present in mitochondria of yeast and animal cells**. In *Methods in Cell Biology* vol. 65: *Mitochondria*. Edited by Wilson L, Matsudaira P. New York: Academic Press; 2001, 463-482.
18. Millar AH, Heazlewood JL: **Genomic and proteomic analysis of mitochondrial carrier proteins in *Arabidopsis***. *Plant Physiol* 2003, **131**:443-453.
19. Lescuyer P, Strub JM, Luche S, Diemer H, Martinez P, Van Dorsselaer A, Lunardi J, Rabilloud T: **Progress in the definition of a reference human mitochondrial proteome**. *Proteomics* 2003, **3**:157-167.
20. Werhahn W, Braun HP: **Biochemical dissection of the mitochondrial proteome from *Arabidopsis thaliana* by three-dimensional gel electrophoresis**. *Electrophoresis* 2002, **23**:640-646.
21. Andersson SG, Karlberg O, Canback B, Kurland CG: **On the origin of mitochondria: a genomics perspective**. *Philos Trans R Soc Lond B Biol Sci* 2003, **358**:165-177; discussion 177-179.
22. Karlberg O, Canback B, Kurland CG, Andersson SG: **The dual origin of the yeast mitochondrial proteome**. *Yeast* 2000, **17**:170-187.
23. Adams KL, Daley DO, Whelan J, Palmer JD: **Genes for two mitochondrial ribosomal proteins in flowering plants are derived from their chloroplast or cytosolic counterparts**. *Plant Cell* 2002, **14**:931-943.
24. Gray MW, Burger G, Lang BF: **The origin and early evolution of mitochondria**. *Genome Biol* 2001, **2**:reviews1018.1-1018.5.
25. Marc P, Margeot A, Devaux F, Blugeon C, Corral-Debrinski M, Jacq C: **Genome-wide analysis of mRNAs targeted to yeast mitochondria**. *EMBO Rep* 2002, **3**:159-164.
26. Andersson SG, Zomorodipour A, Andersson JO, Sicheritz-Ponten T, Alsmark UC, Podowski RM, Naslund AK, Eriksson AS, Winkler HH, Kurland CG: **The genome sequence of *Rickettsia prowazekii* and the origin of mitochondria**. *Nature* 1998, **396**:133-140.
27. Sweetlove LJ, Heazlewood JL, Herald V, Holtzapffel R, Day DA, Leaver CJ, Millar AH: **The impact of oxidative stress on *Arabidopsis* mitochondria**. *Plant J* 2002, **32**:891-904.
28. Taylor SW, Fahy E, Murray J, Capaldi RA, Ghosh SS: **Oxidative post-translational modification of tryptophan residues in cardiac mitochondrial proteins**. *J Biol Chem* 2003, **278**:19587-19590.
29. Hodges PE, McKee AH, Davis BP, Payne WE, Garrels JI: **The Yeast Proteome Database (YPD): a model for the organization and presentation of genome-wide functional data**. *Nucleic Acids Res* 1999, **27**:69-73.
30. Rabilloud T, Kieffer S, Procaccio V, Louwagie M, Courchesne PL, Patterson SD, Martinez P, Garin J, Lunardi J: **Two-dimensional electrophoresis of human placental mitochondria and protein identification by mass spectrometry: toward a human mitochondrial proteome**. *Electrophoresis* 1998, **19**:1006-1014.
31. Heazlewood JL, Whelan J, Millar AH: **The products of the mitochondrial *orf25* and *orfB* genes are F(O) components in the plant F(I)F(O) ATP synthase**. *FEBS Lett* 2003, **540**:201-205.
32. Mewes HW, Frishman D, Guldener U, Mannhaupt G, Mayer K, Mokrejs M, Morgenstern B, Munsterkötter M, Rudd S, Weil B: **MIPS: a database for genomes and protein sequences**. *Nucleic Acids Res* 2002, **30**:31-34.