tion that because selection of the knockout *P. falciparum* lines took place in vitro, failure of processes involved in evading the host immune system (such as the correct display of PfEMP1 on the host erythrocyte surface) would not be detrimental for growth in vitro. Of course, the exportome is not just important for enabling infected erythrocytes to adhere to endothelial cells and to evade immune destruction. Indeed, it has long been appreciated that the parasite must also acquire essential nutrients from the external milieu, which requires the tubulo-vesicular network and the expression of transporter proteins in the erythrocyte membrane (Lauer et al., 1997). The parasite may also need to remodel the structure of the host erythrocyte and increase the stability of the erythrocyte membrane to allow for expansion as the parasite grows (and possibly also to block invasion by new parasites) (Pei et al., 2007). One would predict that essential genes mediating parasite growth and metabolic functions would be conserved across *Plasmodium* species. Indeed, a portion of the *P. falciparum* exportome is conserved in a variety of *Plasmodium* species infecting humans and rodents. However, this "conserved core" of the *Plasmodium* exportome consists of only 10 to 30 components, depending on the prediction algorithms (van Ooij et al., 2008). Given the small number of conserved core genes, why did Maier et al. discover so many *P. falciparum* exportome genes that are essential for parasite survival in vitro? The failure to recover particular mutant parasites may reflect the poor efficiency of *P. falciparum* transfection, but a more interesting possibility is that these "essential" genes reflect more complex metabolic requirements for the intraerythrocytic replication of *P. falciparum* relative to other malaria parasites. Clearly, more work is needed to better define the exportome proteins that are vital for parasite replication within the host erythrocyte and for preventing parasite destruction in vivo. The valuable study by Maier et al. provides a useful toolbox to start obtaining answers to these questions.

REFERENCES

Charpian, S., and Przyborski, J.M. (2008). Traffic *9*, 157–165.

Hiller, N.L., Bhattacharjee, S., van Ooij, C., Liolios, K., Harrison, T., Lopez-Estraño, C., and Haldar, K. (2004). Science *306*, 1934–1937.

Lauer, S.A., Rathod, P.K., Ghori, N., and Haldar, K. (1997). Science *276*, 1122–1125.

Maier, A.G., Rug, M., O'Neill, M.T., Brown, M., Chakravorty, S., Szestak, T., Chesson, J., Wu, Y., Hughes, K., Coppel, R.L., et al. (2008). Cell, this issue.

Marti, M., Good, R.T., Rug, M., Knuepfer, E., and Cowman, A.F. (2004). Science *306*, 1930–1933.

Pei, X., Guo, X., Coppel, R., Bhattacharjee, S., Haldar, K., Gratzer, W., Mohandas, N., and An, X. (2007). Blood *110*, 1036–1042.

Ralph, S.A., and Scherf, A. (2005). Curr. Opin. Microbiol. *8*, 434–440.

Sargeant, T.J., Marti, M., Caler, E., Carlton, J.M., Simpson, K., Speed, T.P., and Cowman, A.F. (2006). Genome Biol. *7*, R12.

van Ooij, C., Tamez, P., Bhattacharjee, S., Hiller, N.L., Harrison, T., Liolios, K., Kooij, T., Ramesar, J., Balu, B., Adams, J., et al. (2008). PLoS Pathog. *4*, e1000084. 10.1371/journal.ppat.1000084.

The Mitochondrial Proteome: From Inventory to Function

Chris Meisinger,[1] Albert Sickmann,[2] and Nikolaus Pfanner[1,*]

[1]Institut für Biochemie und Molekularbiologie, ZBMZ, Universität Freiburg, 79104 Freiburg, Germany

[2]Rudolf-Virchow-Center/DFG Research Center for Experimental Biomedicine, Universität Würzburg, 97078 Würzburg, Germany

*Correspondence: nikolaus.pfanner@biochemie.uni-freiburg.de

DOI 10.1016/j.cell.2008.06.043

Mitochondria are central to cellular energetics, metabolism, and signaling. In this issue, Pagliarini et al. (2008) report the largest compendium of mammalian mitochondrial proteins to date. Together with proteomic studies in yeast, this study represents an important step toward the systematic characterization of the mitochondrial proteome and of mitochondrial diseases.

Mitochondria are cellular powerhouses, synthesizing the bulk of the ATP used by eukaryotic cells. The cellular function of mitochondria, however, is not limited to bioenergetics. They play crucial roles in the metabolism of amino acids and lipids, the biosynthesis of heme and iron-sulfur clusters, cell signaling, and apoptosis. Mitochondrial proteins are encoded by two genomes. In humans, the mitochondrial genome codes for only 13 proteins; the remaining 99% of mitochondrial proteins are encoded by nuclear genes and are synthesized as precursors on cytosolic ribosomes and then are imported into mitochondria (Bolender et al., 2008). To understand the role of mitochondria in health and disease, it is important to know the protein composition of this organelle. In this issue, Pagliarini et al. (2008) now report a major step in compiling a comprehensive compendium of mammalian mitochondrial proteins (the MitoCarta).

Previous studies have led to the identification of close to 700 different mitochondrial proteins in mammals, roughly half of the presumed number (Gaucher et al., 2004; Forner et al., 2006). By comparison, the Pagliarini et al. study establishes a list of nearly 1100 genes coding for mouse mitochondrial proteins by combining experimental identification, bioinformatic analysis, and literature curation (Figure 1). About a quarter of the proteins of the mouse mitochondrial proteome do not have known biological functions; a similar proportion of proteins with unknown functions have been reported in the yeast mitochondrial proteome (Reinders et al., 2006). The Pagliarini et al. study represents an impressive effort, with analysis of the protein content of mitochondria from 14 different mouse tissues. About 750 different mitochondrial proteins were found per tissue, and roughly a third of the proteins were found in all tissues, including the majority of subunits of the oxidative phosphorylation (OXPHOS) machinery. Considering the presumed coverage of the mitochondrial proteome, a given tissue may express 900–1000 different mitochondrial proteins. Remarkably, this number is similar to the number of mitochondrial proteins estimated for yeast (~1000) (Reinders et al., 2006), suggesting that single-celled organisms may have similar demands on the composition and functional diversity of their mitochondria as do the individual tissues of higher eukaryotes. Thus, tissue diversity is a likely reason for the larger size of the mitochondrial proteome in mammals compared to yeast.

Which approaches are likely to lead to the reliable identification of constituents of the mitochondrial proteome? The most direct approach uses mass spectrometry, which requires a highly pure mitochondrial preparation (Sickmann et al., 2003; Gaucher et al., 2004; Prokisch et al., 2004; Reinders et al., 2006) or that the analysis is paired with a quantitative assessment of contaminants by comparing crude and purified mitochondrial fractions and identifying proteins enriched in the purified fraction (Forner et al., 2006; Pagliarini et al., 2008). Pagliarini et al. directly identified ~710 proteins from purified mammalian mitochondria. Only studies of the yeast proteome have identified a greater number of authentic mitochondrial proteins (~850) (Reinders et al., 2006). Protein tagging with green fluorescent protein (GFP) is widely used to assess the intracellular location of proteins. Although genome-wide tagging studies have provided important information on subcellular distribution, the results for individual proteins require critical evaluation. Tags can interfere with the proper import of proteins or even lead to mislocalization (Sickmann et al., 2003). Pagliarini et al. derived only a relatively small number of mitochondrial proteins in their MitoCarta database using a GFP-tagging approach. Moreover, they integrated data from various sources—such as bioinformatic predictions of mitochondrial targeting signals, homology comparisons to previously identified yeast mitochondrial proteins, and literature searches—to obtain the compendium of ~1100 genes for mitochondrial proteins.

Although integrating data from multiple sources is a valuable approach, it is critical to assess the validity of the sources. Prediction of mitochondrial localization is typically based on the presence of an amino-terminal presequence that is found in less than 50% of mitochondrial proteins (Sickmann et al., 2003; Bolender et al., 2008). The quality of protein databases also varies significantly. For example, the *Saccharomyces* genome database sets a high standard for gene and protein annotation, whereas some proteomic studies of higher eukaryotes overestimated the number of identified proteins due to the presence of multiple entries for the same protein.

How complete is the mouse mitochondrial proteome? Based on a reference set of mitochondrial proteins, Pagliarini et al. estimate that their compendium covers more than 85% of genes coding for mitochondrial proteins. Thus, considering the false discovery rate of ~10%, they calculate that mice may contain only ~1130 bona fide genes for mitochondrial proteins. However, their mass spectrometric analysis did not lead to the identification of all known subunits of the OXPHOS machinery, whereas in a study of the yeast mitochondrial proteome (with an estimated coverage of 84%), all of the known OXPHOS subunits were identified (Reinders et al., 2006). The reference set of yeast mitochondrial proteins includes numerous proteins of very low abun-

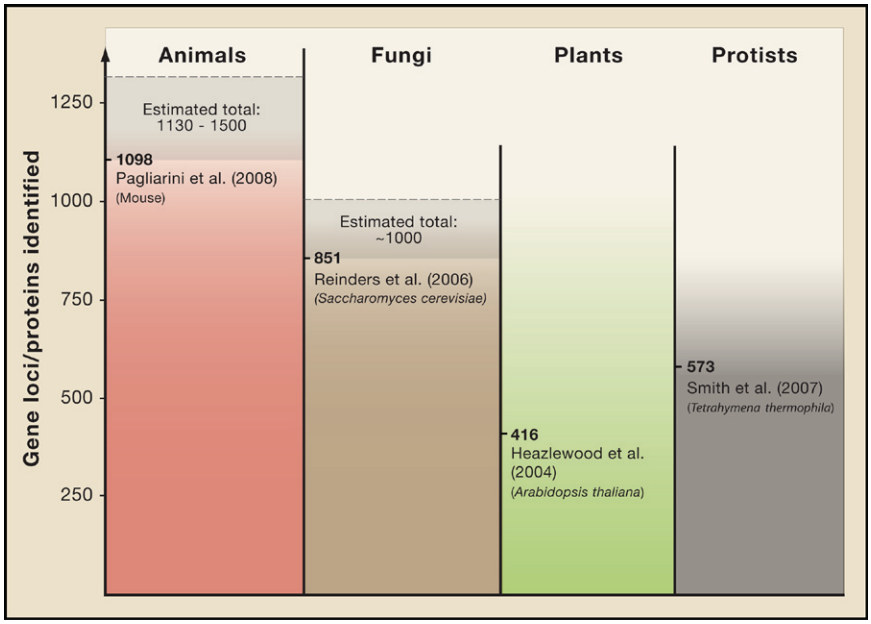

Figure 1. Characterization of Mitochondrial Proteomes

Shown are the number of gene loci/proteins identified in the most recent mitochondrial proteomes in various species (Pagliarini et al., 2008; Reinders et al., 2006; Heazlewood et al., 2004; Smith et al., 2007). All studies analyzed purified mitochondrial fractions by mass spectrometry. In addition to proteomic approaches, Pagliarini et al. (2008) also used bioinformatic analysis and literature curation to obtain a list of mitochondrial proteins, which they have named the MitoCarta.

dance, which were identified by genetic studies, whereas the mouse reference set may contain more proteins of higher abundance. It is worth noting that in the first large-scale proteomic study of yeast mitochondria (~750 proteins), a coverage of 90% was calculated based on the reference set available at that time (Sickmann et al., 2003). However, the subsequent study identified ~850 proteins (PROMITO dataset), and based on an updated reference set, a coverage of 84% was calculated (Reinders et al., 2006) (Figure 1). This places the current estimate of different yeast mitochondrial proteins at ~1000 (the coverage of the first study is now calculated to be ~75%). Although Pagliarini et al. (2008) clearly have performed the most comprehensive study of the mammalian mitochondrial proteome to date, we speculate that the coverage of mitochondrial proteins may be lower than 85% and that the total number of mammalian genes for mitochondrial proteins could approach 1500.

From here, a challenge will be to take this inventory and determine how individual proteins function and are integrated into networks. The genome-wide deletion of genes in yeast and homology searches pave the way toward classifying the function of many new mitochondrial proteins (Sickmann et al., 2003; Prokisch et al., 2004; Bolender et al., 2008). The studies are supported by large-scale protein network analyses in yeast. In mammals, large-scale studies using RNAi will also assist in assigning functions to newly identified mitochondrial proteins. Pagliarini et al. used phylogenetic profiling to identify new factors that are involved in the assembly of complex I of the mitochondrial respiratory chain. Mammalian complex I contains ~45 different subunits that are assembled at the mitochondrial inner membrane in a complicated multistep process (Lazarou et al., 2007). Pagliarini et al. constructed a phylogenetic tree of more than 40 eukaryotes and analyzed the presence of complex I subunits. They noted that complex I, which was present in the bacterial ancestor of mitochondria, was lost at least four times in evolution, for instance in several yeast species. Assuming that species that lost complex I also lost the assembly factors for this complex, then they have identified 19 proteins with the same evolutionary history. They showed experimentally that several of these proteins were required for complex I function. Most importantly, they identified a new gene, *C8orf38*, that is mutated in an inherited human disease of complex I. This demonstrates how the large-scale approach combined with a refined analysis leads to important new findings at the single protein level.

Mitochondria are not independent organelles with little connection to the rest of the cell but rather are integrated into numerous functional, metabolic, and signaling networks with other cellular compartments. Moreover, in both higher and lower eukaryotes, a growing number (10% or more) of mitochondrial proteins possess a dual cellular localization (Sickmann et al., 2003; Pagliarini et al., 2008). Future challenges include analyzing the supramolecular assemblies that form large machines in mitochondrial subcompartments and defining the mitochondrial signaling networks and how they are integrated in the cell.

REFERENCES

Bolender, N., Sickmann, A., Wagner, R., Meisinger, C., and Pfanner, N. (2008). EMBO Rep. *9*, 42–49.

Forner, F., Foster, L.J., Campanaro, S., Valle, G., and Mann, M. (2006). Mol. Cell. Proteomics *5*, 608–619.

Gaucher, S.P., Taylor, S.W., Fahy, E., Zhang, B., Warnock, D.E., Ghosh, S.S., and Gibson, B.W. (2004). J. Proteome Res. *3*, 495–505.

Heazlewood, J.L., Tonti-Filippini, J.S., Gout, A.M., Day, D.A., Whelan, J., and Millar, A.H. (2004). Plant Cell *16*, 241–256.

Lazarou, M., McKenzie, M., Ohtake, A., Thorburn, D.R., and Ryan, M.T. (2007). Mol. Cell. Biol. *27*, 4228–4237.

Pagliarini, D.J., Calvo, S.E., Chang, B., Sheth, S.A., Vafai, S.B., Ong, S.E., Walford, G.A., Sugiana, C., Boneh, A., Chen, W.K., et al. (2008). Cell, this issue.

Prokisch, H., Scharfe, C., Camp, D.G., Xiao, W., David, L., Andreoli, C., Monroe, M.E., Moore, R.J., Gritsenko, M.A., Kozany, C., et al. (2004). PLoS Biol. *2*, e160.

Reinders, J., Zahedi, R.P., Pfanner, N., Meisinger, C., and Sickmann, A. (2006). J. Proteome Res. *5*, 1543–1554.

Sickmann, A., Reinders, J., Wagner, Y., Joppich, C., Zahedi, R., Meyer, H.E., Schönfisch, B., Perschil, I., Chacinska, A., Guiard, B., et al. (2003). Proc. Natl. Acad. Sci. USA *100*, 13207–13212.

Smith, D.G.S., Gawryluk, R.M.R., Spencer, D.F., Pearlman, R.E., Siu, K.W.M., and Gray, M.W. (2007). J. Mol. Biol. *374*, 837–863.