Running head: The potato tuber mitochondrial proteome

Corresponding author:

Dr. Ian Max Møller
Department of Molecular Biology and Genetics
Science and Technology
Aarhus University
Forsøgsvej 1
DK-4200 Slagelse
Denmark

E-mail address: ian.max.moller@agrsci.dk

Research area: Cell Biology
Title: The potato tuber mitochondrial proteome

Salvato, Fernanda\textsuperscript{a,e)}, Havelund, Jesper F.\textsuperscript{b,c)}, Chen, Mingjie\textsuperscript{a)}, Rao, R.S.P.\textsuperscript{a)}, Rogowska-Wrzesinska, Adelina\textsuperscript{c)}, Jensen, Ole N.\textsuperscript{c)}, Gang, David R.\textsuperscript{d)}, Thelen, Jay J.\textsuperscript{a)} and Møller, Ian Max\textsuperscript{b)}

\textsuperscript{a)} Department of Biochemistry and Interdisciplinary Plant Group, University of Missouri-Columbia, Christopher S. Bond Life Sciences Center, Columbia, MO 65211, USA
\textsuperscript{b)} Department of Molecular Biology and Genetics, Science and Technology, Aarhus University, Forsøgsvej 1, DK-4200 Slagelse, Denmark
\textsuperscript{c)} Department of Biochemistry and Molecular Biology, University of Southern Denmark, Campusvej 55, DK-5230 Odense M, Denmark
\textsuperscript{d)} Institute of Biological Chemistry, Washington State University, Pullman, Washington 99164 USA
\textsuperscript{e)} Present address: Institute of Biology, State University of Campinas, São Paulo, Brazil

One-sentence summary:
A high-coverage potato tuber mitochondrial proteome uncovers many new proteins and functions, especially in coenzyme and iron metabolism, and many posttranslational modifications
Footnote:

This study was supported by grants from the Danish Council for Independent Research – Natural Sciences (to IMM) and a 2012 Sabbatical fellowship (to JJT) from the OECD Co-operative Research Programme: Biological Resource Management for Sustainable Agricultural Systems.
Mitochondria are called the powerhouses of the cell. To better understand the role of mitochondria in maintaining and regulating metabolism in storage tissues, highly purified mitochondria were isolated from dormant potato tubers (*Solanum tuberosum* L. cv. Folva) and their proteome investigated. Proteins were resolved by one-dimensional gel electrophoresis, and tryptic peptides were extracted from gel slices and analyzed by liquid chromatography-tandem mass spectrometry using an Orbitrap XL. Using four different search programs, a total of 1060 non-redundant proteins were identified in a quantitative manner using normalized spectral counts including as many as five-fold more “extreme” proteins (low mass, high pI, hydrophobic) than previous mitochondrial proteome studies. We estimate that this compendium of proteins represents a high coverage of the potato tuber mitochondrial proteome (possibly as high as 85%). The dynamic range of protein expression spanned 1800-fold and included nearly all components of the electron transport chain, tricarboxylic acid cycle, and protein import apparatus. Additionally, we identified 71 pentatricopeptide repeat proteins, 29 membrane carriers/transporters, a number of new proteins involved in coenzyme biosynthesis and iron metabolism, the pyruvate dehydrogenase kinase, and a type 2C protein phosphatase that may catalyze the dephosphorylation of the pyruvate dehydrogenase complex. Systematic analysis of prominent post-translational modifications (PTM) revealed that >50% of the identified proteins harbor at least one modification. The most prominently observed class of PTMs was oxidative modifications. This study reveals approximately 500 new or previously unconfirmed plant mitochondrial proteins and outlines a facile strategy for unbiased, near-comprehensive identification of mitochondrial proteins and their modified forms.
Plant mitochondria participate in a number of processes in the plant cell depending on
the cell, tissue, or organ type, the developmental stage, and the environmental
conditions (Rasmusson et al. 2010, Millar et al. 2011). Important examples are energy
metabolism, photorespiration, amino acid biosynthesis, coenzyme (vitamin)
biosynthesis, and programmed cell death. All of these processes require that the
mitochondria can exchange metabolic intermediates and information with the rest of the
cell via membrane carriers and signaling pathways. Plant mitochondria also import
>95% of their proteins across the inner and outer mitochondrial membranes (IMM and
OMM, respectively). Finally, mitochondria are semi-autonomous organelles capable of
growing and dividing, and as such they perform DNA replication, DNA transcription,
and protein biosynthesis, in addition to protein import.

All of these processes require proteins and the composition of the plant mitochondrial
proteome therefore changes depending on the conditions (Millar et al. 2005). It has been
estimated that as many 2000-3000 gene products belong to the mitochondria in
Arabidopsis, but it is not known how many are expressed in a single tissue under a
specific set of conditions (Millar et al. 2005, Millar et al., 2006, Cui et al., 2011). To
date the most complete global proteomic study of isolated mitochondria identified 416
proteins in mitochondria from Arabidopsis cell cultures grown under standard
conditions with no photosynthesis (Heazlewood et al. 2004). In the intervening years a
number of more focused proteomic studies have been performed mainly on Arabidopsis
mitochondria swelling the number of identified plant mitochondrial proteins to 660 non-
redundant proteins (Heazlewood et al., 2004, Duncan et al., 2011, Klodmann et al.
2011, Lee et al., 2012, Tan et al., 2012). Missing from this list are many essential plant
mitochondrial activities including regulatory proteins, transcription factors, metabolite
translocators, and the wealth of tRNA synthases and pentatricopeptide repeat (PPR)
proteins (Havelund et al. 2013).

In yeast (Saccharomyces cerevisiae) about 850 proteins have been identified in isolated
mitochondria, which is thought to provide 85% coverage of the proteome in this species
(Premsl et al. 2009). A total of 1117 proteins were identified in mitochondria from the
nematode Caenorhabditis elegans (Li et al. 2009). The total mammalian mitochondria
is predicted to contain 1050-1400 proteins (Calvo and Mootha 2010) and a global
proteomic study of mitochondria isolated from fourteen mouse tissues using one-dimensional polyacrylamide gel electrophoresis (1D-PAGE) followed by liquid chromatography and two-dimensional mass spectrometry (LC-MS/MS) identified 1098 different gene products (Pagliarini et al. 2008).

Considering that the total plant mitochondrial proteome is thought to contain almost twice as many proteins as the nematode and mammalian mitochondrial proteomes, it is likely that an in-depth proteomic study of one type of plant mitochondria would identify more than 1000 proteins or about twice the previous most complete study. It is likely that newly discovered mitochondrial proteins would be of low abundance, which conceivably would lead to the discovery of new pathways and regulatory proteins, as well as providing additional information about established pathways.

In the present study, we describe the proteome of potato tuber mitochondria (POM). The potato tuber is a large, relatively homogeneous storage organ and potato tuber mitochondria are therefore expected to be similarly homogeneous. There are well-established methods for isolating these mitochondria in a purified, intact and functional form (Neuburger et al. 1982, Struglics et al. 1993, Considine et al. 2003) and their basic properties are therefore well-characterized. In addition, the potato genome has recently been sequenced (Xu et al. 2011), which greatly simplifies the task of identifying potato proteins. Using a prefractionation by preparative SDS-PAGE followed by high-mass accuracy LC-MS/MS we have achieved a high proteome coverage of 1060 proteins, discovered a number of new functions, and identified a large number of posttranslational modifications, particularly oxidative modifications.

RESULTS

The isolated potato tuber mitochondria are highly purified

The isolation of pure organellar fractions from crude source material is the critical phase of any subproteome analysis. It has previously been established that very pure mitochondria can be obtained from potato tubers, which only contain 0.2% peroxisomes or plastids on a protein basis (Neuburger et al. 1982, Struglics et al. 1993, Considine et al. 2003). We employed this established procedure for the isolation of mitochondria with a recognized low contamination level.
We also tested the mitochondrial purity by Western blotting using subcellular marker proteins. The proteins 14-3-3 and anti-enolase were used as reference markers for the cytosolic compartment, while alpha carboxytransferase (α-CT) and pyruvate dehydrogenase E1 component subunit alpha (PDE1-α) were used as markers for the plastidic and the mitochondrial compartment, respectively. Anti-PDE1-α demonstrated the expected mitochondrial enrichment in the mitochondrial fraction, while there was no detectable cytosolic or plastidic contamination (Figure S1).

The potato tuber mitochondrial proteome contains more than 1000 proteins

Previous Arabidopsis mitochondrial proteome studies, both shotgun and targeted, have collectively identified a set of 660 non-redundant proteins (Heazlewood et al., 2004, Duncan et al., 2011, Klodmann et al. 2011, Lee et al., 2012, Tan et al., 2012). These studies employed primarily either two-dimensional (2D) gel electrophoresis or gel-free analyses followed by tandem mass spectrometry (MS). In the present study we performed one-dimensional gel electrophoresis followed by liquid chromatography-tandem mass spectrometry (GeLC-LC-MS) to overcome limitations of 2D gels while allowing for preparative level proteome interrogation. Using GeLC-MS, proteins from each replicate were prefractionated by 12% SDS-PAGE and sectioned into 40 gel slices prior to trypsin digestion, effectively reducing sample complexity prior to LC-MS/MS (Figure 1). To maximize the number of assignments, four complementary database-searching programs were employed simultaneously. Using a false discovery rate (FDR) cutoff of 1% at the protein level, this approach resulted in 7346 assigned non-redundant peptides representing 1060 protein groups including 23 that are mitochondrially encoded. The average coverage was 24.9% (Table I, Table S1).

Figure 2 shows the overlapped peptides and proteins identified using MSGF-DB, ProLuCID, Sequest, and Mascot. A comparison of these different search algorithms revealed that MSGF-DB outperformed the other search engines by identifying 85% of the total peptides, compared to 76% identified by ProLuCID, 46% by Sequest, and 37% by Mascot.

Evaluation of the mitochondria proteome by transient fluorescence assay in tobacco epidermal cell
To evaluate the false positive rate of the mitochondria proteome data, 20 different proteins were selected for further evaluation (Table S1, Figure 3, Figure S2). These selected proteins either were of relatively low abundance and/or their mitochondrial localization was at variance with previous reports. In other words, they were potential contaminants. An EYFP fluorescence tag was in frame fused at their C-terminal; each construct was mixed with a mitochondrial marker (CD3-986) and co-transformed into tobacco leaves for subcellular localization by fluorescence microscope. Out of the total 20 constructs, 19 were detected in tobacco leaves at variable levels, and 18 of them were confirmed to localize to mitochondria. PGSC0003DMP400005581, which encodes a putative inorganic pyrophosphatase, was found to localize to chloroplasts (Figure 3). Its ortholog in *Arabidopsis* (At5g09650) was confirmed to localize to chloroplasts (Schulze et al., 2004). So this protein is likely to be a contaminant.

Two proteins (PGSC0003DMP400035788 and PGSC0003DMP400024128) were found with dual localization. PGSC0003DMP400035788, annotated as pentatricopeptide repeat-containing protein, was found to localize to both mitochondria and chloroplasts (Figure S2). Its ortholog in *Arabidopsis* (At3g46870) was identified as a plastid protein by organellar proteomics (Kleffmann et al., 2004); it also was identified as a Zn$^{2+}$-IMAC interaction protein in mitochondria (Tan et al, 2010). Its dual localization confirmation here resolved the controversy between these two independent investigations (Kleffmann et al, 2004; Tan et al, 2010). PGSC0003DMP400024128, which encodes an RNA-binding protein 24-like, was found to localize to mitochondria and another unknown organelle (Figure S2). This protein is low abundance and none of the subcellular prediction algorithms predicted it to localize to mitochondria.

Two proteins (PGSC0003DMP400012523 and PGSC0003DMP400050514) were observed to localize to mitochondria (Figure S2), and these observations contradict previous reports. PGSC0003DMP400012523’s *Arabidopsis* ortholog (At1g78590) was reported to localize to the cytosol although that conclusion is open to question as no cytosolic marker was used (Chai et al., 2006). PGSC0003DMP400050514’s *Arabidopsis* ortholog (At2g37250) was reported to locate to chloroplasts (Carrari et al., 2005). Since this *Arabidopsis* ortholog falls into the same phylogenetic group as the potato plastidial isoform of adenylate kinase, with which it showed 76% identity, it will require additional phylogenetic analysis to determine their orthologs. The existence of a
mitochondrial isoform of adenylate kinase in potato tuber was first reported by Roberts et al. (1997).

In summary, these EYFP-localization results are consistent with the conclusion that the potato tuber mitochondria are highly purified and that no more than 5% of the 1060 identified mitochondria proteins are false positives. This ratio could be well be overestimated since the candidates for the localization study were purposely selected based on their low abundance and their contradiction with reported results in the literature.

**Categorization of the identified proteins according to properties and function**

Considering proteins only detected in at least two replicates, we have identified 884 protein groups (Table S3), of which 32% were confirmed by ortholog analysis when compared to the Arabidopsis mitochondrial proteome (Heazlewood et al., 2004, Duncan et al., 2011, Klodmann et al. 2011, Lee et al., 2012, Tan et al., 2012). Thus 528 new non-orthologous proteins are presented in the current study, in comparison to all previous Arabidopsis mitochondrial proteome investigations. As shown in Figure 4, the potato mitochondrial proteins were distributed between 4 to 161 kDa in size, between -0.9 to +0.3 in GRAVY score, and with pI values ranging from 4 to 11. A comparison of the physical properties of the potato versus the Arabidopsis mitochondrial proteome (Heazlewood et al., 2004) demonstrates that a much larger number of small proteins with negative GRAVY scores have been identified in the potato mitochondrial proteome (Figure 4).

Assigned potato mitochondrial proteins were annotated by BLAST querying against the non-redundant NCBI database and were classified based on the gene ontology terms related to biological process and molecular function (Figure S3). In addition, each protein was assigned to thirteen general functional categories according to the classification by Heazlewood et al. (2004). The distribution of proteins classified in each category is compared between potato and Arabidopsis mitochondrial proteomes (Table I). The largest functional classes in the potato mitochondrial proteome were energy (12%), metabolism (20%), and protein synthesis (12%). Altogether these proteins represent 45% of total protein abundance based on normalized spectral counting (dNSAF). These proteins are mainly involved in the tricarboxylic acid cycle.
(TCA), oxidative phosphorylation, and protein synthesis (mainly ribosomal and tRNA synthetase proteins). Around 23% of the identified proteins (also, 23% of the abundance) were categorized in the RNA processing, DNA synthesis and processing, transcription or protein fate classes. In addition, approximately 12% of the identified proteins (also 12% of the abundance) were related to transport, or defense/detoxification, or signaling processes.

The number of proteins found in each functional category was higher for potato than for Arabidopsis with the exception of structural organization, although the increase for the energy category was relatively small (1.3 fold) (Table I). The most marked increase in number of proteins identified was seen for RNA processing and protein synthesis where 7.4-fold and 8.4-fold more proteins were found in potato, but also the category transport showed a large increase (3.6 fold). The relative frequency (%) of proteins in each category were very similar, with the exception of the energy category, which was significantly lower in potato, and RNA processing and protein synthesis categories, which were significantly higher. Unclassified proteins, that is, proteins not confidently assigned to any functional class, comprised 21% of the total number of identified proteins (20% of the abundance) in the potato mitochondrial proteome (Table I).

A number of new proteins and enzymes were identified that increase our coverage of known mitochondrial processes and, more interestingly, extend the list of mitochondrial processes for instance with respect to coenzyme and iron metabolism. This will be treated in the Discussion.

**The abundance of mitochondrial proteins varies more than 1800-fold**

It has been shown that in LC-MS/MS based analysis of peptide mixtures raw spectral counts correlate with relative protein abundance between different samples (Liu et al., 2004; Stevenson et al., 2009, Matros et al., 2011). However, to compare the relative abundance of different proteins within the same biological sample spectral counts for each protein must be normalized. One of the most logical and promising strategies for normalization is based on protein length (Zhang et al., 2010).

In the present work, proteins that were detected in at least two replicates (884 proteins) were assigned an abundance factor (Table S3). For this, we calculated the distributed normalized spectral abundance factor (dNSAF) which considers unique spectral counts
and distributes the shared spectral counts matched to peptides assigned to homologous proteins. Spectral counts of each protein were then normalized by protein length (see Methods for details).

Figure 5 shows the log$_{10}$ transformed dNSAFs distribution of the 884 potato mitochondrial proteins. Each bin corresponds to 0.25 order of magnitude difference in the protein abundance. In this way, the potato mitochondrial protein abundance spans three orders of magnitude ranging from -4.82 to -1.57 log$_{10}$ values (Table S3) giving a dynamic range of about 1800 fold as calculated by the normalized spectral count (dNSAF). This is likely a low estimate of dynamic range as the most abundant proteins may be underestimated. The least abundant protein detected in the potato proteome was a ribonuclease (GI 255538392) and the most abundant protein was a phosphate carrier protein (GI 255543593). Considering the previous mitochondrial study in Arabidopsis (Heazlewood et al., 2004), the current study significantly increased coverage of both low- and high-abundance proteins. For example, 53 predicted mitochondrial ribosomal proteins were confidently detected in at least two replicates in the current study with an average abundance of -2.68 [log$_{10}$(dNSAF)] against only three ribosomal proteins detected in the previous Arabidopsis study (Heazlewood et al., 2004). Also, 52 proteins involved in the uptake of a variety of metabolites (ADP/ATP, dicarboxylate, fumarate/succinate, phosphate, malate) including membrane carriers, transporters, and porins were identified in the potato mitochondrial proteome with an average abundance of -3.31, while all previous Arabidopsis studies detected 19 proteins related to the same functional class. Proteins with well-known roles in primary mitochondrial metabolism including proteins involved in oxidative phosphorylation via the respiratory complexes and proteins from the TCA cycle were detected with an average abundance of -2.82 comprising around 28% of total protein abundance. We also quantified low abundance proteins like those involved in RNA processing. For example, we quantified 62 pentatricopeptide repeat (PPR) proteins, which have almost two orders of magnitude difference in expression levels (from -4.67 to -2.69 [log$_{10}$(dNSAF)]) (Table S3).

An overview of the number and abundance of the identified proteins belonging to major mitochondrial metabolic pathways are shown as heat maps in Figures 6 and 7. Figure 6 shows the whole mitochondrion with processes placed in the known or presumptive mitochondrial subcompartment – OMM, intermembrane space (IMS), IMM, matrix – while Figure 7 shows the respiratory complexes and a few associated enzymes located...
in the IMM. Spectral counting should not be used to determine the stoichiometry of subunits in a complex, particularly membrane complexes. Having said that, the apparent wide variation in the amounts of the subunits, e.g. of Complex V, is mainly due to the presence of low-abundance paralogs (Table S3). If we compare the relative abundances of all the F$_1$-ATPase subunits identified, they range between -3.72 and -1.96 [log$_{10}$(dNSAF)] or almost 100-fold (Table S3). However, if we only compare the most abundant paralogs of each of the five subunits identified, the range is much more narrow, between -2.91 and -1.96 [log$_{10}$(dNSAF)], which is about 10-fold or about the same range as the expected 3:3:1:1:1 stoichiometry of the $\alpha$, $\beta$, $\gamma$, $\delta$, $\varepsilon$ subunits (von Ballmoos et al. 2009).

**Prediction programs for mitochondrial localization recognize 63% of all identified proteins**

There are a large number of programs available to predict subcellular localization based on the N-terminal region of proteins that may contain intrinsic targeting peptides (Scheneider and Fechner, 2004). The protein sequences of the identified mitochondrial proteome were analyzed by five different prediction programs: TargetP (Emanuelsson et al., 2000), Predotar version 1.03 (Small et al. 2004), MitoProtII (Claros and Vincens, 1996), iPSORT (Bannai et al., 2002) and WoLF PSORT (Horton et al., 2007). Each program employs different methods for subcellular prediction and usually produces relatively large non-overlapping sets of results (Heazlewood et al., 2004). Each prediction tool returned a variable percentage of mitochondrial positives ranging from approximately 20% (WoLF PSORT) to 52% (MitoProtII) of the potato proteome (Table S4). We also performed a relational evaluation of the overlapping positive predictions by different combinations of the prediction tools. When we considered multiple prediction tools together, the number of positive mitochondrial predictions decreased (Table S4), indicating a much smaller agreement between programs. The combination of MitoProt II and iPSORT provided the best pair for prediction resulting in 37% positive potato mitochondrial predicted proteins (Table S4).

In all, 671 out of 1060 proteins (63% of the total protein set) were predicted to be mitochondrial by at least one program, while 48% were predicted by two or more programs (Fig. 8). From the list of 389 proteins not predicted to be mitochondrial by any program, 44 proteins were confirmed OMM or IMM proteins related to the protein
import system and uptake of metabolites (carriers, transporters, channels, and translocases) which are characterized by the absence of N-terminal targeting presequences (Millar and Heazlewood, 2003; Huang et al., 2010). In addition, 31 well-known mitochondrial proteins involved in the TCA cycle and the respiratory complexes, as well as 11 proteins encoded by mitochondrial genes, were also negative for mitochondrial localization prediction. These results show the limitations in assigning mitochondrial localization based solely on bioinformatic tools. The choice of prediction program can benefit different classes of proteins and the absence of targeting presequences, due to genuine lack of those N-terminal targeting signaling sequences (e.g. OMM or IMM proteins or mitochondrially encoded proteins), or due to incorrect annotation of the N-terminus of the sequences, can result in erroneous interpretation.

In Figure 8, the subcellular prediction information is superimposed on the relative protein abundance. The data reveal that prediction results are comparable irrespective of protein abundance, suggesting that low abundance proteins must also be considered as authentic mitochondrial proteins.

Proteins with uncleaved N-terminal sequences are sometimes further processed

Semi-tryptic database querying resulted in 37 matches to the protein N-terminus indicating that these proteins do not contain a cleavable targeting peptide. Some of these peptides had the terminal Met removed while others did not. The removal of the Met can be related to the length of the side chain of the second amino acid in the sequence (Gliglione and Meinnel, 2001). If the second amino acid is Ala, Gly, Pro, Ser or Thr, the initiating Met is removed by a Met aminopeptidase. On the other hand, if the side-chain is large, in case of Arg, Asn, Asp, Glu, Ile, Leu or Lys, the Met is retained. Only Val appears to have an intermediate-side-chain specificity for Met cleavage, showing the initiating Met cleaved or uncleaved. As shown in Table S5 our results confirm this specificity. The Met was removed when the second side chain was Lys, Ala, Pro, Ser, Thr or Val and not removed when it was Lys, Asn or Asp or Gln. Surprisingly, when the second residue was Gly, some sequences had Met removed while others did not.

Many of the proteins identified with an uncleaved N-terminus were OMM proteins, including porins, or components of the respiratory complexes present in the IMM, such as NADH dehydrogenase, bc, complex, and ATPase. Additionally, two mitochondrially
encoded proteins were identified with intact N-termini, a ribosomal protein (STmpRH_41) and an NADH dehydrogenase subunit (STmpRH_8).

The majority of proteins in the potato mitochondrial proteome are post-translationally modified

The potato mitochondrial proteome was extensively searched for different types of post-translational modifications (PTMs) (Tables S6 and S7). We identified 3,066 PTM sites in a total of 556 proteins which means that 52% of the entire potato mitochondrial proteome determined here exist in modified versions. Individual proteins contained up to eight types of PTMs and up to 50 different modification sites. The proteins with most modifications - aconitase, 2-oxoglutarate dehydrogenase, succinate dehydrogenase and glycine dehydrogenase - are all part of the TCA cycle or associated with it (Table S7).

Mitochondria are not only the sites of oxygen consumption but also one of the sources of cellular reactive oxygen species (ROS) (Møller 2001). We therefore specifically focused our analyses on protein oxidation in the potato mitochondrial proteome. Oxidative modifications of Arg, Pro, Lys, Met, Thr, and Trp were detected. A total of 505 proteins contained at least one type of oxidative modification giving a total of 2,471 sites of modification. As shown in Figure 9, Met sulfoxidation followed by Pro oxidation (either hydroxylation or carbonylation) and Lys hydroxylation are the most abundant types of PTM in the proteome. Most of the modified proteins participate in the TCA cycle, respiratory complexes, protein import assembly, and protein synthesis (ribosomal proteins) (Table S7). The high frequency of Met sulfoxidation is in accordance with several reports that found sulfur amino acids to be more sensitive to oxidation by ROS than other amino acids (Stadtman et al., 2005).

Carbonylated proteins include both high and low abundance proteins spanning a variety of metabolic pathways (energy, metabolism, protein and DNA synthesis, protein fate and transport). Among the highly abundant carbonylated proteins (abundance from -2.99 to -1.56 log10[dNSAF]), up to 50 oxidation sites were mapped to a single protein, glycine dehydrogenase (Table S7). Other examples were proteins involved in the TCA cycle (aconitase, 2-oxoglutarate dehydrogenase, succinate dehydrogenase, citrate synthase, isocitrate dehydrogenase), respiratory complexes (ATPase, cytochrome c1, cytochrome c oxidase), chaperonins, carrier proteins (phosphate, oxoglutarate,
ADP/ATP), pyruvate dehydrogenase complex (E1 α and β subunits) and redox metabolism (peroxiredoxin, Mn superoxide dismutase, monodehydroascorbate reductase). Many low-abundance proteins (\(-4.65\) to \(-3.0 \log_{10}[dNSAF]\)) also showed a large number of oxidation sites, up to 25 sites. Proteins with more than ten oxidation sites include FtsH proteases, monoxygenase, ATPase and branched-chain α-keto acid dehydrogenase (Table S7).

We also detected Asn deamidation (188 proteins), Lys acetylation (31 proteins) and methylation (63 proteins), and Ser/Thr phosphorylation (19 proteins). These proteins showed a wide range of relative abundance (\(-4.56\) to \(-1.56 \log_{10}[dNSAF]\)) and comprised components of TCA cycle and associated enzymes, chaperone system, respiratory complexes, protein import system, RNA processing and carriers proteins.

**DISCUSSION**

The discord between the number of predicted and reported plant mitochondrial proteins is strikingly high – probably more so than for any other organelle. Despite this discordance, no shotgun proteomic study of plant mitochondria has been published since 2005. Many advances in proteomics have occurred during the intervening time, including the development of best practices for unbiased protein prefractionation, sensitive and higher-accuracy mass spectrometers, as well as various refinements in mass spectral data querying. In this study, we leverage developments on each of these fronts to offer a quantitative, unbiased, and comprehensive proteome resource for the plant mitochondrial community, including about 500 new experimentally identified plant mitochondrial proteins. We estimate that the coverage of the mitochondrial proteome is high (possibly as high as 85%) (see Supplementary Discussion). Since the theoretical size of the plant mitochondrial proteome is about twice the size of the the potato tuber proteome, it is likely that the specialized and physiological condition of the tissue and its mitochondria placed a limit on the number of proteins identified. Thus, we can discuss not only proteins found, but also comment with some confidence on proteins or protein groups underrepresented. This will be a recurrent theme in the following discussion as well as in Supplementary Discussion.

**Prediction programs**
A comparison of five different subcellular localization prediction programs, TargetP, Predotar, MitoProtII, iPSORT and WoLF PSORT), revealed a variable number of mitochondrial predictions, ranging from 20 to 52% of the entire experimentally determined mitochondrial proteome (Fig. 8, Table S4). These values are low compared to the previous values of 70 to 90% reported in the original articles for these programs (Emanuelsson et al., 2000; Small et al. 2004, Claros and Vincens, 1996). This could be the result of bias in the protein training sets used in developing the programs. Also, we found that the combination of multiple prediction programs showed low numbers of overlapping positive predictions, indicating the divergence of the prediction methods employed by each tool. Overall, 63% of the potato mitochondrial proteome was predicted to be mitochondrially localized by at least one of the five prediction tools (Figure 8, Table S4). However, authentic mitochondrial proteins involved in the TCA cycle and the respiratory complexes as well as proteins characterized by the absence of targeting presequences (OMM or IMM and mitochondrially encoded proteins) were erroneously predicted to be non-mitochondrial. The number of such false negatives was only partially offset by the use of different algorithms, indicating that the results returned by such tools should be inspected carefully. False positives appear to be less of a problem with these prediction programs. Nevertheless, even if a protein is not predicted to be mitochondrial, one should not assume it is a contaminant as the accuracy of these programs is both variable and limited. As most of these programs are decades old, the development of new improved programs, based on more comprehensive genomic sequences and proteomics data, is required to overcome the lack of accuracy observed.

### Basic metabolic processes

The coverage of the well-known mitochondrial functions - TCA cycle, respiratory chain, transmembrane transporters, and turnover of amino acid, proteins, lipids and fatty acids - was high as shown in Figures 6 and 7 and listed in Table S1. A detailed discussion of the potato tuber proteins identified for each group as well as for glycolysis and signaling is found in Supplementary Discussion. In the following we will discuss a number of interesting protein groups linked to other important mitochondrial functions.

### Biosynthesis of coenzymes and pyrimidines
Vitamins are coenzymes, which humans are unable to synthesize but acquire through their food, mostly from plants. Mitochondria are involved in the biosynthesis of several of the coenzymes in plant cells and we have found a number of key enzymes in these processes, some of which have previously only been predicted.

Although NAD$^+$ is the main pyridine coenzyme in mitochondria, a number of enzymes use NADP$^+$ (Møller and Rasmusson 1998). NAD$^+$ uptake into plant mitochondria was discovered in the 1980’s using POM (Neuburger et al. 1985), but NADP$^+$ uptake has also been inferred at least for pea leaf mitochondria (Bykova and Møller 2001). We here identify an NAD$^+$ carrier as well as an NADH kinase (Table S1). Three NAD(H) kinases are found in Arabidopsis and one of them, NDK3, which appears to prefer NADH as substrate over NAD$, has been reported to be localized to peroxisomes in higher plants while its analogue in yeast is mitochondrial (Waller et al. 2010). The POM NADH kinase is predicted to be mitochondrial by three prediction programs and is of relatively low abundance (~3.87 [log$_{10}$(dNSAF)]) (Tables S1 and S3, Fig. 6). Thus, in order to generate the NADP$^+$/NADPH pool inside the POM, it appears that NAD$^+$ is taken up, reduced by one of the many matrix dehydrogenases, e.g. in the TCA cycle, and then phosphorylated to give NADPH. This does not imply that NADPH/NADP$^+$ ratios in the mitochondria are determined by this pathway, rather that generation of the overall NADP$^+$/NADPH pool size would be governed by the activities of the NAD$^+$ transporter and the NADH kinase.

L-Galactono-1,4-lactone dehydrogenase, the last enzyme in one of the ascorbate biosynthetic pathways, which donates electrons to Complex IV on the outer surface of the IMM (Bartoli et al. 2000), but is physically associated with Complex I (Pineau et al. 2008), was found at above median quantities. Five different enzymes using ascorbic acid mainly for ROS detoxification were also found (see later). Clearly ascorbate must be imported in some form, but the mitochondrial carrier protein family does not appear to include a carrier for ascorbate or dehydroascorbate (Taylor et al. 2010, Palmieri et al. 2011) and no such carrier was found in our study.

Tetrahydrofolate (vitamin B9) and its derivatives catalyze the addition or removal of C1-units, e.g., in the glycine decarboxylate reaction in the mitochondrial matrix where a methyl group is transferred from one glycine molecule to another to form serine. Folates consist of a pterin moiety, a $p$-aminobenzoate moiety and a (poly)glutamate tail and the
three parts are assembled in the mitochondria (Blancquaert et al. 2010). In the POM, we have found three enzymes involved in folate biosynthesis, dihydropterin pyrophosphokinase-dihydropteroate synthase, a molybdopterin biosynthesis protein and polypolyglutamate synthase, four enzymes involved in the interconversion and transfer of different C1-units, 5-formyltetrahydrofolate cycligase, formyltetrahydrofolate deformylase-like, methenyltetrahydrofolate synthase domain-containing protein-like, 5-
methyltetrahydrodoproteinylglutamate-homocysteine methyltransferase-like and finally a folate carrier responsible for exporting folate to make the coenzyme available to the rest of the cell.

Biotin (vitamin B8) is a coenzyme involved in carboxylations and the last steps in its biosynthesis are mitochondrial (Alban 2011). We found the S-adenosylmethionine carrier responsible for the import of one of the precursors, but not predicted to be mitochondrial by any of the prediction programs. We also detected the adrenodoxin reductase involved in biotin biosynthesis and predicted to be mitochondrial by three programs (Table S1). The adrenodoxin reductase uses adrenodoxin to reduce S-S bridges and we find five ferredoxin analogs, which may well be adrenodoxin. Thus, a significant part of the biotin biosynthesis pathway is expressed in POM.

Pyrimidine biosynthesis mostly takes place in the cytosol and the nucleus in mammalian cells, but the fourth reaction in the pathway, catalyzed by dihydroorotate dehydrogenase, takes place on the outer IMM surface (Löffler et al. 2005). Heazlewood et al. (2004) found this enzyme in Arabidopsis mitochondria and in POM it is present at higher than median relative abundance (Table S3) confirming that it is mitochondrial and probably donates its electrons to ubiquinone as in mammalian mitochondria (Fig. 7).

The identification in this one investigation of so many enzymes involved in coenzyme biosynthesis in the mitochondrion is a good demonstration of the power of applying high resolution and highly sensitive mass spectrometry-based proteomic tools to important questions in plant biochemistry.

**DNA and RNA**

We identified DNA polymerases, RNA polymerases, RNA helicase, histones, an histone-modifying enzyme, a topoisomerase, transcription factors and a transcription
termination factor, most at low abundances (Table S1 and S3) and representing almost all of the components required for DNA replication and transcription. Three of the five histones identified are of a type recently reported to be present in plant mitochondria (Zanin et al. 2010).

**Pentatricopeptide repeat (PPR) proteins**

PPR proteins are one of the most prolific protein families in plants while it is virtually absent in animals (Small and Peeters 2000). The majority of the 450 PPR proteins in Arabidopsis are predicted to be mitochondrial where about half are thought to be involved in RNA editing and the remainder in other types of RNA processing (Fujii and Small 2011). In Arabidopsis mitochondria, Heazlewood et al. (2004) found 10 PPR proteins while in potato tuber mitochondria we found 71. From the total of 71 PPR proteins, we relatively quantified (dNSAF) 62 PPR proteins found in at least two replicates. The abundance of these PPR proteins varied by almost two orders of magnitude based on dNSAF values. Only 7 PPR proteins can be considered to be present at high relative abundance (-2.9 to -2.6 \[\log_{10}(\text{dNSAF})\]), whereas 23 can be considered medium relative abundance (-3.9 to -3.0 \[\log_{10}(\text{dNSAF})\]) and 32 low relative abundance (-4.67 to -4.0 \[\log_{10}(\text{dNSAF})\]).

Since 450 PPR proteins are present in Arabidopsis (Fujii and Small 2011) and since the potato tuber genome contains many more genes than the Arabidopsis genome, it is very likely that at least 300 PPR proteins are targeted to the mitochondria in some potato tissue under some condition. In spite of the fact that all proteins encoded in mtDNA are subunits in the constitutively expressed respiratory complexes or appear to be essential housekeeping proteins, the finding of “only” 71 PPR proteins indicates that many RNA-related processes are inactive in POM. In mitochondria from perennial ryegrass reproductive tissues, a transcript profile showed that the expression of several mitochondrially expressed ribosomal proteins was very low (Islam et al. 2013). This may mean that their expression, and the associated RNA-related processes including PPR proteins, is dependent on the tissue and/or environmental conditions.

**Turnover of Reactive Oxygen Species**

The mitochondrion is one of the major sites of ROS production in the cell (Maxwell et al. 1999, Møller 2001, Foyer and Noctor 2003) and it therefore also contains a number
of enzymes or enzyme systems capable of removing ROS. We found three of the four enzymes in the ascorbate-glutathione pathway, glutathione reductase, ascorbate peroxidase, and monodehydroascorbate reductase (dehydroascorbate reductase was not detected) (Jimenez et al. 1998, Chew et al. 2003) with very similar relative abundances (-3.9 to -3.6 \( \log_{10}(dNSAF) \)), except for one monodehydroascorbate reductase (GI 350536875) which presented higher relative abundance (-2.8 \( \log_{10}(dNSAF) \)) (Fig. 5).

Also, we have identified both the classical mitochondrial Mn-SOD as well as a Cu,Zn-SOD, which at least in yeast is found in the mitochondrial intermembrane space (O’Brien et al. 2004). Mn- and Cu,Zn SODs showed high (-1.7 \( \log_{10}(dNSAF) \)) and medium (-3.5 \( \log_{10}(dNSAF) \)) relative abundances, respectively.

We found thioredoxin (-3.1 \( \log_{10}(dNSAF) \)) and thioredoxin reductase, phospholipid hydroperoxide glutathione peroxidase (-3.2 \( \log_{10}(dNSAF) \)) and peroxiredoxin (-2.6 \( \log_{10}(dNSAF) \)) – all in medium relative abundance. In other words, all the four predicted NADPH-consuming enzyme systems involved in removing H\(_2\)O\(_2\) or H\(_2\)O\(_2\)-induced peroxidation products (Møller 2007) are expressed in POM, which appear well armed to deal with oxidative stress.

Finally, we detected two catalases (one at low level – 4.4 and one at medium level (-2.6 \( \log_{10}(dNSAF) \)). Although POM only contain a maximum of 0.2% peroxisomes on a protein basis, the catalases could be contaminants, since catalase makes up approximately 50% of all the peroxisomal protein in potato tubers (Struglics et al. 1993). However, catalase has been reported to be present inside rat heart mitochondria (Radi et al. 1991) so the parsimonious explanation of catalase inside plant mitochondria cannot be excluded.

**Metal ion turnover**

Mitochondria contain many proteins, particularly redox active enzymes, which bind metal ions often at their catalytic site. As mitochondria are also involved in the biosynthesis of FeS centers and hemes, the concentration of metal ion is high in mitochondria (Tan et al. 2010). However, because free metal ions can interact with hydrogen peroxide and form the highly reactive and damaging hydroxyl radical, which can lead to protein oxidation (see below), it is very important for the cell to keep metal ions safely bound to proteins at all stages of metal ion uptake and metabolism (Kell 2009, 2010, Møller et al. 2011). We found a number of prominent metal-containing
proteins including the FeS-proteins and the Cu/Fe-containing cytochromes in the ETC, but also Mn- and Cu/Zn-SOD and TCA cycle enzymes like aconitase (Fe). In addition, we found the Fe-storing ferritin (low abundance) and the complete pathway for the biosynthesis of FeS clusters – frataxin, iron-sulfur cluster assembly proteins, iron-sulfur cluster scaffold protein, iron-sulfur cluster co-chaperone protein, cysteine desulfurase and ferredoxin (all medium abundance) (Ye and Rouault 2010).

A majority of potato mitochondrial proteins are post-translationally modified

More than half the proteins detected in the potato mitochondrial proteome were post-translationally modified on at least one site. And about 100 proteins had more than ten PTM sites. The most modified protein, glycine dehydrogenase, had as many as 50 PTM events of six different kinds including 24 Met oxidations, 3 Asp deamidations and 12 Pro oxidations (Table S7). In most cases the spectral counts were very low for each modified peptide identified. It is likely that the coverage of modified peptides were underestimated as PTMs such as Lys or Arg oxidation causes missed trypsin cleavage sites resulting in larger peptides that are more difficult to ionize and fragment.

The large number of PTM sites on some proteins could give rise to a very large number of differentially modified versions of the protein unless the PTMs are introduced in an ordered rather than a random manner (Thelen and Miernyk, 2012). A relatively unexplored research area is the interaction between different PTMs. At least in one case we know that such an interaction takes place – The oxidation of a Met residue to Met-sulfoxide next to a potential phosphorylation site abolishes phosphorylation on that site and this is likely a regulatory mechanism as this first step in Met oxidation is reversible (Huang et al., 2010).

In Supplementary Discussion we analyze phosphorylated proteins, oxidized proteins and proteins with acetylated and methylated lysines.

Conclusions

A comprehensive shotgun proteomic investigation of potato tuber mitochondria led to the detection of 1060 proteins, which appear to represent more than 85% of the proteins expressed in the POM. A significant disparity was observed between the proteins actually identified in the proteome and those predicted to be targeted to the proteome by the five most commonly used prediction algorithms. Less than 50% of the proteins
determined to be present in the mitochondrial proteome were predicted to be thus localized by at least two of these programs. Indeed, only 63% were thus predicted by at least one program. This clearly demonstrates that current bioinformatics tools, while often useful, can be very inadequate at predicting biological function and localization.

The broad coverage observed in the POM proteome allowed for investigation of processes that are important for mitochondrial function, such as metabolism and respiration and post-translational regulation of protein function. The ETC was well covered in the POM proteome, with all complexes being well represented. In addition, three different alternative NAD(P)H dehydrogenases were present at significant levels, as were all members of the TCA cycle, except fumarase, suggesting that non-cyclic modes of operation are likely to be functional. In addition, the biosynthesis of many coenzymes was well represented in the proteome, supporting the role of the mitochondrion in production of these important molecules. The general processes of amino acid biosynthesis and lipid metabolism were well supported in the proteome. The fact that a number of glycolytic enzymes were detected can be explained by their association with the OMM. A number of other processes important to mitochondrial function and replication were well supported, including transport, protein synthesis, regulation of gene expression within the mitochondria, and metal ion regulation.

Finally, a large number of protein PTMs were found in the mitochondrial proteome, with over 50% of proteins possessing such modifications and several proteins possessing over a dozen modifications, such as phosphorylation and oxidative modifications. The ease with which these modifications were found is a testament to the power of the approach used. The fact that so many proteins contained at least one, and many contained more than one PTM, suggests that this mode of functional regulation plays an important role in the mitochondria. It further suggests that approaches beyond transcriptional profiling will be not only desired, but required, if we are to be able to better understand how mitochondria help plants respond to stress and changing environments.

**MATERIALS AND METHODS**

**Isolation of mitochondria**
Mitochondria were isolated from potato (*Solanum tuberosum* L., cv. Folva) tubers by the method of Considine et al. (2003) with minor modifications. All procedures were done at 4°C. Peeled tubers were homogenized using a juice extractor into an equal volume of extraction medium (0.6 M mannitol, 20 mM MOPS-KOH (pH 7.3), 2 mM EDTA, 25 mM cysteine and 0.3 % (w/v) BSA). The homogenate was left standing for 5 min allowing starch to sediment before it was filtered through 2 layers of cotton towel and centrifuged at 3000 g for 5 min. The supernatant was centrifuged at 18000 g for 10 min to recover an organelle pellet. The pellet was resuspended in mannitol buffer (0.3 M mannitol, 10 mM MOPS-KOH (pH 7.3), 0.1 % (w/v) BSA) and layered on top of a Percoll step gradient of 50%, 28% and 20% in mannitol buffer. After centrifugation at 40000 g for 30 min the pale mitochondrial band was recovered from the 28%/50% Percoll interface. After washing, the mitochondria were further purified on a second self-forming Percoll gradient of 28% Percoll in sucrose buffer (0.3 M sucrose, 10 mM MOPS-KOH (pH 7.3), 0.1 % (w/v) BSA). The protein concentration was estimated by measuring absorbance at 280 nm using a NanoDrop 1000 Spectrophotometer (Thermo Scientific).

Gel electrophoresis

One-dimensional sodium dodecyl sulfate polyacrylamide gel electrophoresis (1D-SDS–PAGE, 20 × 20 cm) performed according to (Laemmli 1970) was employed as a pre-fractionation step. Isolated mitochondria (250 µg) and size marker proteins (protein ladder, Thermo Scientific # 26630) were dissolved in sample buffer (NuPAGE®), placed at 65°C for 15 min and resolved by 12% SDS-PAGE. After protein migration, proteins were stained with colloidal Coomassie Brilliant Blue. Then, each gel lane was sliced in 40 segments and diced in approximately 1 mm cubes followed by gel destaining and in gel digestion as described by Balbuena et al. (2011).

Liquid chromatography-mass spectrometry

A two column system was used. Both columns were in-house packed with C18 reverse phase material (ReproSil, C18 AQ 3 µm (Dr. Maisch, Germany)) in fused silica. The pre-column had an inner diameter of 100 µm and a length of 2 cm. The analytical column had an inner diameter of 75 µm and a length of 15 cm. The flow of 250 nl/min was delivered by an EASY nLC system (Thermo Fisher Scientific, Odense, Denmark).
and the peptides were eluted using a 21 min gradient from 0% to 40% B buffer (A
buffer: 0.1% formic acid, B buffer: 0.1% formic acid/90% acetonitrile, all v/v). The
flow from the analytical column was coupled to an LTQ-Orbitrap XL mass
spectrometer (Thermo Fisher Scientific). The instrument was run in positive ion mode.
Full scan MS spectra (from m/z 300-1800) were acquired in the Orbitrap with resolution
r = 60,000 at m/z 400. The up to 5 most intense ions with charge states larger than +1
were sequentially isolated for fragmentation in the linear ion trap using collision-
induced dissociation. Former target ions selected for fragmentation were dynamically
excluded for 30 s. For accurate mass measurements the lock mass option was enabled in
MS mode and the polydimethylcyclosiloxane (PCM) ions generated in the electrospray
process (m/z = 445.120025) were used for internal recalibration in real time, resulting in
a mass accuracy of approximately 2 ppm.

**Database searching and protein identification**

Acquired MS/MS spectra were searched using different algorithms against the potato
protein database (http://www.potatogenome.net). To estimate the protein false discovery
rate (FDR), randomized sequences were combined with the forward database in a
concatenated format. MASCOT server 2.3.01, SEQUEST, MS-GFDB (Kim et al.,
2010) and ProLuCID (Xu et al., 2006) were used to interpret the data set using similar
parameters. MASCOT and SEQUEST were integrated within the Proteome Discoverer
1.3 software package (Thermo Scientific). Raw files (Thermo) were converted to ms2
and mzXML files prior to analysis by ProLuCID and MS-GFDB. All search engines
were configured to the following parameters: MW range between 200 and 2,000; 1,000
ppm of precursor ion tolerance; fragment tolerance 1.0 Da oxidation of methionine and
asparagine deamidation as variable modifications and carbamidomethylation of cysteine
as a static modification; two allowed missed trypsin cleavages. After searches, the
PSMs were filtered with a 5 ppm precursor ion tolerance, achieving a false discovery
rate (FDR) lower than 1% at the protein level for each biological replicate.

All the MS data from each gel slice was searched and merged for each replicate. The
results from different programs were combined in order to complement observations
that came from different processes of data mining, without compromising the accuracy.
To do this, a high stringency on precursors (5 ppm) was applied.
Filtered data were uploaded to the Protein Herder module within the Compass package (Wenger et al., 2011) for protein grouping. At this step, all sets of indistinguishable proteins, which were identified by the same peptides, were combined into protein groups.

**Purity assessment of mitochondrial preparations**

Total protein and mitochondrial protein-enriched fractions from potato tubers were separated by SDS-PAGE. From each fraction, 20 μg of proteins was loaded on the gel. Blotting was performed as previously described by Stevenson et al. (2009) to PVDF membranes (GE Healthcare Ltd, UK). Antibodies directed against 14-3-3, enolase, plastidic alpha carboxytransferase (α-CT) and mitochondrial pyruvate dehydrogenase E1 component subunit alpha (PDE1-α) were used to assess mitochondrial protein enrichment in the preparations. Immunodetection of proteins bound to PVDF membranes was performed using the colorimetric alkaline phosphatase and peroxidase substrate detection system (Sigma-Aldrich).

**Subcellular Localization Study**

Total RNA was isolated from potato tuber by using the method described by Kumar et al., (2007). A cDNA library was then synthesized by M-MLV reverse transcriptase reaction (Promega, WI, US) with random hexamer. Potato cDNA was PCR amplified from cDNA library by a pair of gene-specific primers; BamHI and Not I restriction enzyme cutting sites were added to forward and reverse primers, respectively (Table S8). The PCR fragments were cloned into pGEM-T easy vector (Promega, WI, US), the insertions were dropped off by BamHI and NotI double digestion, and the insertions were cloned into pE6c entry vector, such that an EYFP fluorescence tag was fused in frame with potato proteins at the C-terminus (Dubin et al., 2008). The cassette was then moved into the binary vector pSITE-0B by Gateway LR clonase (Chakrabarty et al., 2007).

For the co-localization assay, mitochondria marker (CD3-986) binary plasmid with CFP fluorescence tag was ordered from the Arabidopsis Biological Resource Center (Nelson et al., 2007). The binary constructs were transformed into Agrobacterium AGL1 strain. The potato EYFP fusion construct was individually mixed with the CFP mitochondrial marker and co-transformed into tobacco leaves. Transformation was performed by
following the procedure described by Sparkes et al. (2006). An Olympus IX70 microscope controlled by MetaMorph (Version 6.3 v6, Molecular Device) software was used for image capture, and Chroma filter 49001 and 49003 were applied for the CFP and EYFP channel, respectively.

**Post-translational modifications**

Multiple post-translational modifications (PTM) were searched against a concatenated database containing forward and randomized protein sequences previously identified in the general proteome searching approach. The searches were performed using SEQUEST, MASCOT and MS-GFDB programs. For each round of PTM searching a maximum of two different variable modifications (PTMs) were configured. Precursor ion tolerance was set to 100 ppm and 2 missed tryptic cleavages were allowed. Carbamidomethylation of cysteine was configured as a static modification and the PTMs of interest were configured as variable modification. After searches, the data set was filtered using 5 ppm as precursor ion tolerance, achieving FDR lower than 1%. Only PTMs detected in at least 2 replicates were considered. The spectra assigned to peptides with oxidation sites were searched against the database allowing all types of oxidation sites at the same time to make sure that the oxidation sites were reproducible compared to the first database search.

**Semi-tryptic database searching**

All MS data was reanalyzed using MS-GFDB program against the identified protein sequences database allowing semi-tryptic peptides and 100 ppm of precursor ion tolerance. Carbamidomethylation of cysteine was configured as a static modification and methionine oxidation as variable modification. The PSMs obtained were filtered out using 5 ppm as precursor ion tolerance, resulting in a FDR lower than 1%.

**Prediction of Subcellular localization**

Predictions of subcellular localization for identified proteins were taken using the full-length protein sequences and five programs namely TargetP version 1.1 (http://www.cbs.dtu.dk/services/TargetP/, Emanuelsson et al., 2000), Predotar version 1.03 (http://urgi.versailles.inra.fr/predotar/predotar.html, Small et al. 2004), MitoProtII (http://ihg.gsf.de/ihg/mitoprot.html, Claros and Vincens, 1996), iPSORT
Molecular weight and isoelectric point

Molecular weight and isoelectric point of proteins were obtained using Compute pI/Mw web server (http://web.expasy.org/compute_pi/, Bjellqvist et al. 1994, Gasteiger et al. 2005). The grand average hydropathy (GRAVY) score was computed based on Kyte-Doolittle hydropathy scale (Kyte and Doolittle 1982). Bioinformatics computations were performed using Python ver. 2.7. Arabidopsis mitochondrial proteome list is based on Supplementary Table 1 in Heazlewood et al. (2004).

Protein abundance calculation

Total number of PSM (peptide spectrum matched) for each protein identified was extracted from Protein Herder module of Compass software (Wenger et al., 2011) for each biological replicate. Only proteins identified in at least two replicates were considered for quantification based on the distributed normalized spectral abundance factor approach (dNSAF, Zhang et al., 2010). Spectral counts (SpC) for each protein or group of proteins were used to estimate protein abundance.

For protein abundance estimation, the PSMs produced by each program were unified by an in-house developed script. Only different scans for a given protein detected by multiple programs were considered for calculation of total spectral counts, which means that repeated scans were considered only once in the calculation of total spectral counts. Therefore, the total spectra counts of each protein was calculated by the sum of PSMs (non-redundant scans) produced by different algorithms.

The dNSAF considers the spectral counts of shared peptides and distributes them based on a distribution factor ($d$):

$$d_k = \frac{uSpC_k}{\sum_{n} uSpC_n}$$

where $k$ denotes a protein identity and $n$ is the total number of proteins, while $uSpC$ represents the sum of spectra matched to peptides mapping uniquely to the respective
protein. So, for a given protein $k$, $d$ is equal to the total $uSpC$ mapped to this protein divided by the total number of $uSpC$ from all $n$ proteins with protein $k$ shares peptide(s).

Then, the dNSAF for a given protein, $k$, is calculated as follows:

$$dNSAF_k = \frac{\left[ uSpC + d \times sSpC \right] \times L_k}{\sum_{n} \left[ uSpC + d \times sSpC \right] \times L_n}$$

where $sSpC$ corresponds to the total number of spectra from peptides shared among related proteins and $L$ represents the length of the protein. In this way, the shared spectral counts ($sSpC$) are distributed amongst the proteins, avoiding that the same spectrum will be counted multiple times.

**Annotation and functional classification**

Protein annotation and classification were performed based on matching the protein sequences against the non-redundant NCBI database and Gene Ontology functional classification using the Blast2GO tool (Conesa et al., 2005). Protein sequences were aligned using the BLASTP algorithm against nr NCBI database using the following parameters: report a maximum of 5 blast hits, 0.1 for the expected value and minimum high scoring segment pairs (HSPs) length equal to 33. Mapping and annotation steps were performed using Blast2GO default values. GOslim annotation was performed using the plant slim. GO distribution graphs related to biological process and molecular function were generated and analyzed at the fourth level of depth. Also, each protein was assigned to general functional categories according to Heazlewood et al. (2004).

**Acknowledgements**

JFH is grateful to Drs. Katherine Beard and Lee J. Sweetlove for a study visit to their laboratory at University of Oxford.

**Authorship**

JFH, JJT and IMM designed the research, JFH, MC and JJT performed the experiments, FS, JFH, JJT, RSRR, and IMM contributed new computational tools and analyzed the data, and FS, JFH, ARW, ONJ, DRG, JJT and IMM wrote the paper.
Supplemental Data

Supplementary Discussion

Figure S1: SDS-PAGE and corresponding immunoblots of enriched mitochondrial protein fraction and total protein extract from potato tubers. A, Commassie blue-stained SDS-PAGE gel loaded with 20 µg of protein per lane. Positions of protein standards are shown on the left (in kilodaltons). B, protein blots of mitochondrial enriched fraction and total protein extract (fractions in A). The molecular masses of the protein bands are shown on the left (in kilodaltons).

Figure S2: Subcellular localization study on a selected subset of proteins identified. See Figure 2 and Table S2 for details.

Figure S3: GO term distribution related to biological process and molecular function of the proteins identified in the potato tuber mitochondria.

Table S1: List of 1060 protein groups identified by GeLC-MS/MS in isolated mitochondria from Solanum tuberosum tubers.

Table S2: List of selected candidate proteins for validation by subcellular localization in tobacco leaf.

Table S3: Label-free quantification based on spectral counting (884 proteins).

Table S4: Mitochondrial localization prediction results (for potato and Arabidopsis mitochondrial proteomes) with different combinations of prediction programs.

Table S5: Proteins with non-cleaveable N-terminal presequences identified by LC-MS/MS.

Table S6: Monoisotopic mass changes used to identify the post-translational modifications.

Table S7: Posttranslational modifications detected in the mitochondrial proteome of Solanum tuberosum.
Repository data

Raw files as well peptide and protein identification files exported for each search engine (SEQUEST, MASCOT MSGF-DB and ProLuCID) have been deposited in a public repository - the ProteomeXchange Consortium (http://proteomecentral.proteomexchange.org) via the PRIDE partner repository (Vizcaino et al. 2013) with the dataset identifier PXD000149 available at http://www.ebi.ac.uk/pride/. Phosphopeptides and respective spectra data have been deposited at P3DB (Plant Protein Phosphorylation Database) available at http://p3db.org/PotatoMitochondriaPhosData.php.

LITERATURE CITED


Day DA, Hanson JB (1977) Effect of Phosphate and Uncouplers on Substrate Transport and Oxidation by Isolated Corn Mitochondria. Plant Physiol 59: 139-144


Kell DB (2009) Iron behaving badly: inappropriate iron chelation as a major contributor to the aetiology of vascular and other progressive inflammatory and degenerative diseases. BMC Med Genom 2: 2

Kell DB (2010) Towards a unifying, systems biology understanding of large-scale cellular death and destruction caused by poorly liganded iron: Parkinson's, Huntington's, Alzheimer's, prions, bactericides, chemical toxicology and others as examples. Arch Toxicol 84: 825-889


Melo AMP, Roberts TH, Møller IM (1996) Evidence for the presence of two rotenone-
insensitive NAD(P)H dehydrogenases on the inner surface of the inner membrane of
dr potato tuber mitochondria. Biochim Biophys Acta 1276: 133-139

Michalecka AM, Svensson ÅS, Johansson FI, Agius SC, Johanson U, Brennicke A,
Binder S, Rasmusson AG (2003) Arabidopsis genes encoding mitochondrial type II
NAD(P)H dehydrogenases have different evolutionary origin and show distinct

Milenkovic D, Kozjak V, Wiedemann N, Lohaus C, Meyer HE, Guiard B, Pfanner N,
Meisinger, C (2004) Sam35 of the mitochondrial protein sorting and assembly
machinery is a peripheral outer membrane protein essential for cell viability. J Biol
Chem 279(21): 22781-22785


mitochondrial proteome. Trends Plant Sci 10: 36-43

mitochondria and plastids. Current Opinion in Plant Biology 9: 610-615

Millar AH, Whelan J, Soole KL, Day DA (2011) Organization and regulation of
mitochondrial respiration in plants. Annu Rev Plant Biol 62: 79-104

Pyrroline-5-Carboxylate-Proline Cycle in Plants by Uncoupled Expression of Proline

Trends Plant Sci. 3: 21-27


**FIGURE LEGENDS**

Figure 1: Workflow chart for the isolation of mitochondria from potato tubers followed by protein separation, identification and spectral counting-based quantification. First, mitochondria were isolated from potato tubers using Percoll gradients. Then, mitochondrial proteins were extracted and fractionated by SDS-PAGE. After gel segmentation, in gel trypsin digestion was performed and peptides were injected into an LC system coupled to a mass spectrometer. Database searching was carried out using four different algorithms (MSGF-DB, ProLuCID, Mascot, Sequest) and the potato genome database (http://www.potatogenome.net). The number of assigned MS/MS spectra for each identified protein was determined and the distributed normalized spectral abundance factor was calculated for protein abundance estimation.

Figure 2: The number of peptides (A) and proteins (B) identified using different database search engines with a <1% protein FDR identification. The total number of non-redundant peptides and proteins were 7346 and 1060, respectively.

Figure 3: Subcellular localization study on a selected subset of proteins identified in this study. Twenty proteins that were identified in purified potato tuber mitochondria (Table S1) were in-frame fused with EYFP fluorescence tag, then transiently co-expressed with the mitochondrial marker CD3-986 in tobacco epidermal cells. The left lane shows the EYFP fluorescence of six out of the 19 constructs where expression was detected. The center lane shows the fluorescence of the mitochondrial marker. The right lane shows an overlay between the two types of fluorescence and green colour indicates coincidence of the two probes, i.e., mitochondrial localization. The remaining 13 constructs are shown in Figure S2. Size bar, 12.5 μm.

Figure 4: Properties of the potato mitochondrial proteome (n = 1060) as compared to the *Arabidopsis* mitochondrial proteome (n = 416; [Heazlewood et al. 2004]). (A) Distribution of molecular mass (kDa), (B) isoelectric point and (C) Grand Average Hydropathy (GRAVY) score. A disproportionately large number of low-molecular-weight proteins with negative GRAVY scores were identified in the potato mitochondrial proteome.
Figure 5: Abundance distribution of the potato mitochondrial proteome. Mitochondrial localization prediction results are superimposed on the protein abundance (spectral counting) [\log_{10}(dNSAF)]. Plus signs (+) indicate mitochondrial localization by as many as five prediction programs and minus sign (-) indicates localization by none. Each category on the x-axis corresponds to 0.25 order of magnitude difference in the abundance of proteins. The difference between the most abundant and the least abundant proteins was 1778 fold. The abundance calculation for all proteins is found in Table S3.

Figure 6: Schematic representation showing proteins that were identified by GeLC-MS-MS in the mitochondrial preparation of potato tubers. Squares represent protein identification and abundance based on spectral counting. Each color represents a range of \log_{10} transformed dNSAF (bottom left). Numbers in parentheses represent number of proteins found within the same range of abundance. Respiratory chain complexes (OXPHOS) are shown in detail in Figure 6. Black text: metabolic intermediates; grey text: enzymes or proteins. Abbreviations: ABC, ABC transporter; AC, aconitase; ACOT, acyl-coenzyme A thioesterase 9; ACS, Long chain acyl-CoA synthetase 9APX, ascorbate peroxidase; AOX, alternative oxidase; ASC, ascorbate; AlaAT, alanine aminotransferase; AspAT, aspartate aminotransferase; BCKDH, branched chain alpha-keto acid dehydrogenase; CAC, carnitine/acylcarnitine carrier protein; CI-CV, respiratory chain complexes (see Figure 7); CoA DHase, isovaleryl-CoA dehydrogenase; CS, citrate synthase; DiT1, oxoglutarate malate transporter; DHA, dehydroascorbate; DHAR, dehydroascorbate reductase; DIC, Mitochondrial dicarboxylate carrier protein; D-LDH, D-lactate dehydrogenase; E1α, pyruvate dehydrogenase subunit α; E1β, pyruvate dehydrogenase subunit β; E2, dihydrolipoamide S-acetyltransferase; E3, Dihydrolipoamide dehydrogenase; ECR, trans-2-enoyl-CoA reductase; EH, Enoyl-CoA hydratase; ER, enoyl-acyl-carrier-protein reductase; ETF, electron transfer flavoprotein; FCL, 5-formyltetrahydrofolate cycloligase; FDF, formyltetrahydrofolate deformylase-like; FUM, fumarase; GABA, γ-aminobutyric acid; GABA-T, GABA transaminase; GDC, glycine decarboxylase complex; GDH, glutamate dehydrogenase; Gly-II, hydroxyacylglutathione hydrolase 3; Gly-I, lactoylglutathione lyase; GPX, glutathione peroxidase; GR, glutathione reductase; GRX, glutaredoxin; GSH; reduced glutathione; GSSG; oxidized glutathione; GST, glutathione-s-transferase; HAD, hydroxyacyl-thioester dehydratase; HMG lyase,
hydroxymethylglutaryl-CoA lyase; ICDH, isocitrate dehydrogenase; IscA, iron sulfur cluster assembly protein; MCAT, acyl-carrier-protein S-malonyltransferase; MCC, 3-methylcrotonyl-CoA carboxylase; MCD, malonyl CoA decarboxylase; MCP, mitochondrial carrier proteins; MG hydratase, methylglutaconyl-CoA hydratase; MDH, malate dehydrogenase; MDHA, monodehydroascorbate; MDHAR, monodehydroascorbate reductase; MFT, mitochondrial folate transporter/carrier-like; MnSOD, superoxide dismutase; MPP, mitochondrial processing peptidase; NAD-ME, NAD-dependent malic enzyme; NDT, nicotinamide adenine dinucleotide transporter; 2-OGDH, 2-oxoglutarate dehydrogenase; PDK, pyruvate dehydrogenase kinase; PGP, phospholipid hydroperoxide glutathione peroxidase; PHB, prohibitin; Pic, phosphate carrier protein; PRDX, peroxiredoxin; PSP, presequence protease; SAM, S-adenosylmethionine; SAM-T, S-adenosylmethionine-dependent methyltransferase; SAMC, S-adenosylmethionine carrier protein; SDH, succinate dehydrogenase; SCaMC, calcium-binding mitochondrial carrier protein; SFC, Succinate/fumarate mitochondrial transporter; SHMT, serine hydroxymethyltransferase; SSA, succinic semialdehyde; SSAD, SSA dehydrogenase; succCoA-syn, succinyl CoA synthetase; ThPP, thiamine pyrophosphate; TIMs, import inner membrane translocases; TOMs, outer membrane translocases; TPC, thiamine pyrophosphate carrier; tRNAs, tRNA synthetases TXNRD, thioredoxin reductase; TX-SS, thioredoxin oxidized; TX-SH, thioredoxin reduced; UCP, uncoupling protein; VDAC, voltage-dependent anion channel

Figure 7: Schematic representation showing the proteins identified by GeLC-MS/MS related to the functionality of mitochondrial respiratory chain complexes. Abbreviations: AOX, alternative oxidase; ATPase, ATP synthase; CI, complex I; CII, complex II; CIII; complex III; CIV, complex IV; CV, complex V; Cyt b, cytochrome b; Cyt c, cytochrome c; Cyt c1, cytochrome c1; COX, cytochrome oxidase; Ddh, dihydroorotate dehydrogenase; ETF, electron transfer flavoprotein; exNDH, external NADH-ubiquinone oxidoreductase; FeS, succinate dehydrogenase iron sulfur subunit; GAPdh, glyceraldehydes-3-phosphate dehydrogenase, GLDH, L-galactono-1,4-lactone dehydrogenase; NAD-DH, NADH dehydrogenase [ubiquinone] subcomplex; Succ DH, succinate dehydrogenase; UCP, uncoupling protein; UQox, ubiquinone oxidized; UQred, ubiquinone reduced

Figure 8: Mitochondrial localization prediction (by iPSORT, MitoProt II, Predotar, TargetP and WoLF PSORT). Predicted to be mitochondrial by one (+), two (++).
(+++), four (++++) or all five (+++++) programs. Note that about 35% of the mitochondrial proteome (both potato and Arabidopsis) are not predicted to be mitochondrial by any of the five prediction programs. Combination details of the predictions are given in Supplemental Table S4.

Figure 9: Post-translational modifications (PTMs) in the potato mitochondrial proteome. The list of PTM analyzed can be found in Table II. Of the 1060 potato mitochondrial protein groups, 559 (53%) had at least one PTM. Methionine oxidation followed by proline and lysine oxidation are the most abundant types of PTM. Individual proteins contained up to nine types of PTMs (Table S6).
Table I: Functional distribution of the mitochondrial proteins from potato and Arabidopsis based on the functional classification made by Heazlewood et al. (2004). The relative frequency (%) of proteins classified in each functional category in the Arabidopsis and potato proteome are shown. The fold change in protein numbers (potato/Arabidopsis) are in brackets. The frequency of the relative protein abundance (% dNSAF) of each functional category in the potato proteome is also shown.

<table>
<thead>
<tr>
<th>Function</th>
<th>Arabidopsis</th>
<th>Potato</th>
<th>% dNSAF</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td># proteins</td>
<td>%</td>
<td># proteins</td>
</tr>
<tr>
<td>Energy</td>
<td>98</td>
<td>23.6</td>
<td>125</td>
</tr>
<tr>
<td>Metabolism</td>
<td>81</td>
<td>19.5</td>
<td>216</td>
</tr>
<tr>
<td>Protein fate</td>
<td>53</td>
<td>12.7</td>
<td>102</td>
</tr>
<tr>
<td>DNA synthesis and processing</td>
<td>9</td>
<td>2.2</td>
<td>23</td>
</tr>
<tr>
<td>Transcription</td>
<td>8</td>
<td>1.9</td>
<td>13</td>
</tr>
<tr>
<td>RNA processing</td>
<td>14</td>
<td>3.4</td>
<td>104</td>
</tr>
<tr>
<td>Protein synthesis</td>
<td>15</td>
<td>3.6</td>
<td>126</td>
</tr>
<tr>
<td>Defense, stress, detoxification</td>
<td>16</td>
<td>3.8</td>
<td>30</td>
</tr>
<tr>
<td>Communication/signaling</td>
<td>19</td>
<td>4.6</td>
<td>39</td>
</tr>
<tr>
<td>Transport</td>
<td>19</td>
<td>4.6</td>
<td>59</td>
</tr>
<tr>
<td>Miscellaneous</td>
<td>4</td>
<td>1.0</td>
<td>0</td>
</tr>
<tr>
<td>Structural organization</td>
<td>9</td>
<td>2.2</td>
<td>5</td>
</tr>
<tr>
<td>Unknown</td>
<td>71</td>
<td>17.1</td>
<td>218</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>416</td>
<td></td>
<td>1060</td>
</tr>
</tbody>
</table>
**Mitochondrial proteome analysis**

1. **Isolation of mitochondria**

   - Percoll gradients

2. **Protein extraction and SDS-PAGE fractionation**

   - 40 slices are excised from each gel lane

3. **In gel digestion**

4. **LC-MS/MS analysis**

   - LTQ Orbitrap XL Thermo mass spectrometer
   - Mass spectra collected

5. **Database searching**

   - MSGF-DB
   - SEQUEST
   - PROLUCID
   - MASCOT
   - Combination of results from different search engines

6. **Spectral counting quantification**

   \[
   (dNSAF)_k = \frac{(SAF)_k}{\sum (SAF)_n}
   \]
Figure 2: Number of peptides (A) and proteins (B) identified using different database searching engines <1% protein FDR identification. The total number of non-redundant peptides and proteins were 7346 and 1060, respectively.
Figure 3
Figure 4: Properties of the potato mitochondrial proteome (n = 1060) as compared to the *Arabidopsis* mitochondrial proteome (n = 426; [Heazlewood et al. 2004]). (A) Distribution of molecular mass (kDa), (B) isoelectric point and (C) Grand Average Hydropathy (GRAVY) score. A disproportionately large number of low-molecular-weight proteins with negative Gravy Scores were identified in the potato mitochondrial proteome.)
Figure 5: Mitochondrial localization prediction results are superimposed on the protein abundance (spectral counting) $[\log_{10}(\text{dNSAF})]$. Plus signs (+) indicate mitochondrial localization by as many prediction programs and minus sign (-) indicates localization by none. Each category on the x-axis corresponds to 0.25 order of magnitude difference in the abundance of proteins. The difference between the most abundant and the least abundant proteins was 1778 fold.
Figure 6: Schematic representation showing proteins that were identified by GeLC-MS-MS in the mitochondrial preparation of potato tubers. Squares represent protein identification and abundance based on spectral counting. Each color represents a range of Log10 transformed dNSAF (bottom left legend). Numbers in parentheses represent number of proteins found within the same range of abundance. Respiratory chain complexes (OXPHOS) are showed in detail in Figure 8. Abbreviations: ABC, ABC transporter; AC, aconitase; ACOT, acyl-coenzyme A thioesterase 9; ACS, Long chain acyl-CoA synthetase 9APX, ascorbate peroxidase; AOX, alternative oxidase; ASC; ascorbate; AlaAT, alanine aminotransferase; AspAT, aspartate aminotransferase; BCKDH, branched chain alpha-keto acid dehydrogenase; CAC, carnitine/acylcarnitine carrier protein; CI-CV, respiratory chain complexes (see Figure 7); CoA DHase, isovaleryl-CoA dehydrogenase; CS, citrate synthase; DiT1, oxoglutarate malate transporter; DHA, dehydroascorbate; DHAR, dehydroascorbate reductase; DIC, Mitochondrial dicarboxylate carrier protein; D-LDH, D-lactate dehydrogenase; E1α, pyruvate dehydrogenase subunit α; E1β, pyruvate dehydrogenase subunit β; E2, dihydrolipoamide S-acetyltransferase; E3, Dihydrolipoamide dehydrogenase; ECR, trans-2-enoyl-CoA reductase; EH, Enoyl-CoA hydratase; ER, enoyl-acyl-carrier-protein reductase; ETF, electron transfer flavoprotein; FCL, 5-formyltetrahydrofolate cycloligase; FDF, formyltetrahydrofolate deformylase-like; FUM, fumarase; GABA, γ-aminobutyric acid; GABA-T, GABA transaminase; GDC, glycine decarboxylase complex; GDH, glutamate dehydrogenase; Gly-II, hydroxyacylglutathione hydrolase 3; Gly-I, lactoylglutathione lyase; GPX, glutathione peroxidase; GR, glutathione reductase; GRX, glutaredoxin; GSH; reduced glutathione; GSSG; oxidized glutathione; GST, glutathione-s-transferase; HAD, hydroxacyl-thioester dehydratase; HMG lyase, hydroxymethylglutaryl-CoA lyase; ICDH, isocitrate dehydrogenase; IscA, iron sulfur cluster assembly protein; MCAT, acyl-carrier-protein S-malonyltransferase; MCC, 3-methylcrotonyl-CoA carboxylase; MCD, malonyl CoA decarboxylase; MCP, mitochondrial carrier proteins; MG hydratase, methylglutaconyl-CoA hydratase; MDH, malate dehydrogenase; MDHA, monodehydroascorbate; MDHAR, monodehydroascorbate reductase; MFT, mitochondrial folate transporter/carrier-like; MnSOD, superoxide dismutase; MPP, mitochondrial processing peptidase; NAD-ME, NAD-dependent malic enzyme; NDT, nicotinamide adenine dinucleotide transporter; 2-OGDH, 2-oxoglutarate dehydrogenase; PDK, pyruvate dehydrogenase kinase; PGP, phospholipid hydroperoxide glutathione peroxidase; PHB, prohibitin; Pic, phosphate carrier protein; PRDX, peroxiredoxin; PSP, presequence protease; SAM, S-adenosylmethionine; SAM-T, S-adenosylmethionine-dependent methyltransferase; SAMC, S-adenosylmethionine carrier protein; SDH, succinate dehydrogenase; SCaMC, calcium-binding mitochondrial carrier protein; SFC, Succinate/fumarate mitochondrial transporter; SHMT, serine hydroxymethyltransferase; SSA, succinic semialdehyde; SSAD, SSA dehydrogenase; succCoA-syn, succinyl CoA synthetase; ThPP, thiamine pyrophosphate; TIMs, import inner membrane translocases; TOMs, outer membrane translocases; TPC, thiamine pyrophosphate carrier; tRNAs, tRNA synthetases TXNRD, thioredoxin reductase; TX-SS, thioredoxin oxidized; TX-SH, thioredoxin reduced; UCP, uncoupling protein; VDAC, voltage-dependent anion channel.
Figure 7: Schematic representation showing the proteins identified by GeLC-MS/MS related to the functionality of mitochondrial respiratory chain complex.

Abbreviations: AOX, alternative oxidase; ATPase, ATP synthase; CI, complex I; CII, complex II; CIII; complex III; CIV, complex IV; CV, complex V; Cyt b, cytochrome b; Cyt c, cytochrome c; Cyt c1, cytochrome c1; COX, cytochrome oxidase; Ddh, dihydroorotate dehydrogenase; ETF, electron transfer flavoprotein; exNDH, external NADH-ubiquinone oxidoreductase; FeS, succinate dehydrogenase iron sulfur subunit ; GAPdh, Glyceraldehyde 3-phosphate dehydrogenase, GLDH, L-galactono-1,4-lactone dehydrogenase; NAD-DH, NADH dehydrogenase [ubiquinone] subcomplex; Succ DH, succinate dehydrogenase; UCP, uncoupling protein; UQox, ubiquinone oxidized; UQred, ubiquinone reduced.
Figure 8: Mitochondrial localization prediction (by iPSORT, MitoProt II, Predotar, TargetP and WoLF PSORT). Predicted to be mitochondrial by one (+), two (++), three (+++), four (++++) or all five (+++++) programs. Note that about 35% of the mitochondrial proteome (both potato and Arabidopsis) are not predicted to be mitochondrial by any of the five prediction programs. Combination details of the predictions are given in the Supplemental TableS3.
Figure 9: Post-translational modifications (PTMs) in the potato mitochondrial proteome. The list of PTM analyzed can be found in Table 3. Of the 1060 potato mitochondrial protein groups, 559 (53%) had at least one PTM. Methionine oxidation followed by proline and lysine oxidation are the most abundant types of PTM. Individual proteins contained up to nine types of PTMs (Table S6)