Cell Metabolism, Volume *33*

Supplemental information

Quantitative high-confidence human mitochondrial proteome and its dynamics in cellular context

Marcel Morgenstern, Christian D. Peikert, Philipp Lübbert, Ida Suppanz, Cinzia Klemm, Oliver Alka, Conny Steiert, Nataliia Naumenko, Alexander Schendzielorz, Laura Melchionda, Wignand W.D. Mühlhäuser, Bettina Knapp, Jakob D. Busch, Sebastian B. Stiller, Stefan Dannenmaier, Caroline Lindau, Mariya Licheva, Christopher Eickhorst, Riccardo Galbusera, Ralf M. Zerbes, Michael T. Ryan, Claudine Kraft, Vera Kozjak-Pavlovic, Friedel Drepper, Sven Dennerlein, Silke Oeljeklaus, Nikolaus Pfanner, Nils Wiedemann, and Bettina Warscheid

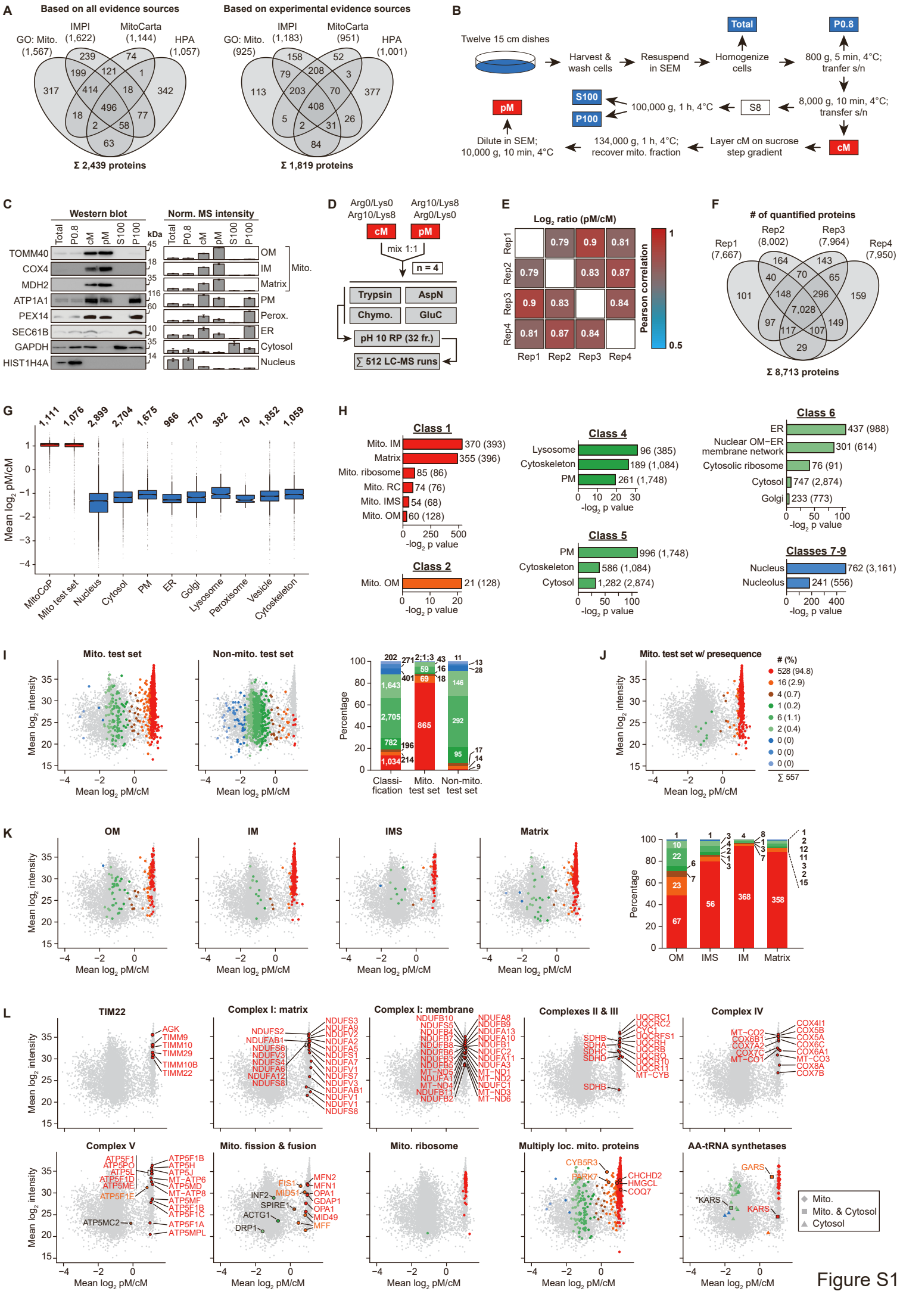
A
Based on all evidence sources
GO: Mito. (1,567)
IMPI (1,622)
MitoCarta (1,144)
HPA (1,057)
239
74
199
121
1
317
414
18
342
496
18
77
2
58
63
Σ 2,439 proteins
Based on experimental evidence sources
GO: Mito. (925)
IMPI (1,183)
MitoCarta (951)
HPA (1,001)
158
52
79
208
3
113
203
70
377
408
5
2
31
26
84
Σ 1,819 proteins
B
Twelve 15 cm dishes
Harvest & wash cells
Resuspend in SEM
Homogenize cells
Total
800 g, 5 min, 4°C; tranfer s/n
P0.8
8,000 g, 10 min, 4°C; transfer s/n
S8
100,000 g, 1 h, 4°C
S100
P100
cM
Layer cM on sucrose step gradient
134,000 g, 1 h, 4°C; recover mito. fraction
pM
Dilute in SEM; 10,000 g, 10 min, 4°C
C
Western blot
Norm. MS intensity
Total
P0.8
cM
pM
S100
P100
kDa
TOMM40
COX4
MDH2
ATP1A1
PEX14
SEC61B
GAPDH
HIST1H4A
OM
IM
Matrix
PM
Perox.
ER
Cytosol
Nucleus
Mito.
D
Arg0/Lys0
Arg10/Lys8
Arg10/Lys8
Arg0/Lys0
cM
pM
mix 1:1
n = 4
Trypsin
AspN
Chymo.
GluC
pH 10 RP (32 fr.)
∑ 512 LC-MS runs
E
Log2 ratio (pM/cM)
Rep1
Rep2
Rep3
Rep4
0.79
0.9
0.81
0.79
0.83
0.87
0.9
0.83
0.84
0.81
0.87
0.84
Pearson correlation
1
0.5
F
# of quantified proteins
Rep1 (7,667)
Rep2 (8,002)
Rep3 (7,964)
Rep4 (7,950)
164
143
40
70
65
101
148
296
159
7,028
97
117
107
149
29
Σ 8,713 proteins
G
1,111
1,076
2,899
2,704
1,675
966
770
382
70
1,852
1,059
Mean log2 pM/cM
MitoCoP
Mito test set
Nucleus
Cytosol
PM
ER
Golgi
Lysosome
Peroxisome
Vesicle
Cytoskeleton
H
Class 1
Mito. IM
370 (393)
Matrix
355 (396)
Mito. ribosome
85 (86)
Mito. RC
74 (76)
Mito. IMS
54 (68)
Mito. OM
60 (128)
-log2 p value
Class 2
Mito. OM
21 (128)
Class 4
Lysosome
96 (385)
Cytoskeleton
189 (1,084)
PM
261 (1,748)
Class 5
PM
996 (1,748)
Cytoskeleton
586 (1,084)
Cytosol
1,282 (2,874)
Class 6
ER
437 (988)
Nuclear OM–ER membrane network
301 (614)
Cytosolic ribosome
76 (91)
Cytosol
747 (2,874)
Golgi
233 (773)
Classes 7-9
Nucleus
762 (3,161)
Nucleolus
241 (556)
I
Mito. test set
Non-mito. test set
Mean log2 intensity
Mean log2 pM/cM
Percentage
Classi-fication
Mito. test set
Non-mito. test set
J
Mito. test set w/ presequence
(%)
528 (94.8)
16 (2.9)
4 (0.7)
1 (0.2)
6 (1.1)
2 (0.4)
0 (0)
0 (0)
0 (0)
Σ 557
K
OM
IM
IMS
Matrix
Mean log2 intensity
Mean log2 pM/cM
Percentage
OM
IMS
IM
Matrix
L
TIM22
AGK
TIMM9
TIMM10
TIMM29
TIMM10B
TIMM22
Complex I: matrix
NDUFS3
NDUFA9
NDUFV2
NDUFA2
NDUFA5
NDUFS1
NDUFA7
NDUFV1
NDUFS7
NDUFV3
NDUFAB1
NDUFV1
NDUFS8
NDUFS2
NDUFAB1
NDUFS6
NDUFV3
NDUFS4
NDUFA6
NDUFA12
NDUFS8
Complex I: membrane
NDUFA8
NDUFB9
NDUFA13
NDUFA10
NDUFB1
NDUFC2
NDUFA11
NDUFA3
MT-ND1
MT-ND2
NDUFC1
MT-ND3
MT-ND6
NDUFB10
NDUFS5
NDUFB4
NDUFB7
NDUFB8
NDUFB6
NDUFB3
NDUFB5
MT-ND5
NDUFA1
MT-ND4
NDUFB11
NDUFB2
Complexes II & III
UQCRC1
UQCRC2
CYC1
UQCRFS1
UQCRH
UQCRB
UQCRQ
UQCR10
UQCR11
MT-CYB
SDHB
SDHA
SDHC
SDHD
SDHB
Complex IV
COX4I1
COX5B
COX5A
COX6C
COX6A1
MT-CO3
COX8A
COX7B
MT-CO2
COX6B1
COX7A2
COX7C
MT-CO1
Complex V
ATP5F1
ATP5PO
ATP5L
ATP5D
ATP5ME
ATP5F1E
ATP5MC2
ATP5F1B
ATP5H
ATP5J
MT-ATP6
ATP5MD
MT-ATP8
ATP5MF
ATP5F1B
ATP5F1C
ATP5F1A
ATP5MPL
Mito. fission & fusion
MFN2
MFN1
OPA1
GDAP1
OPA1
MID49
MFF
FIS1
MID51
INF2
SPIRE1
ACTG1
DRP1
Mito. ribosome
Multiply loc. mito. proteins
CYB5R3
PARK7
CHCHD2
HMGCL
COQ7
AA-tRNA synthetases
GARS
KARS
*KARS
Mito.
Mito. & Cytosol
Cytosol

Figure S1

Figure S1. Assessment of the Subcellular Fractionation Workflow and of Data Acquired in Subtractive Proteomics Experiments

Related to Figure 1

(A) Venn diagram representing unique proteins with mitochondrial annotation in different data repositories based on all entries (left) and entries with experimental evidence for mitochondrial protein localization (right). GO: Mito, GO-CC term 'mitochondrion'; IMPI, integrated mitochondrial protein index; HPA, human protein atlas.

(B) Outline of the subcellular fractionation protocol established in this study for the preparation of crude and gradient-purified mitochondrial fractions (highlighted in red) from human cell lines (HEK293T, HeLa). Samples highlighted in blue were used for spatial proteomics experiments. SEM, sucrose/EDTA/MOPS buffer; s/n, supernatant; P, pellet; S, supernatant; cM/pM, crude/pure mitochondria.

(C) Abundance profiles of organellar marker proteins across different subcellular fractions established by Western blot (left) and MS analysis (right). Error bars indicate SEM for n = 3 and the range for n = 2. Norm., normalized; Total, cell lysate; P0.8, nuclear fraction; cM/pM, crude/pure mitochondria; S100/P100, cytosolic/microsomal fraction; OM/IM, outer/inner mitochondrial membrane; PM, plasma membrane; Perox., peroxisome; ER, endoplasmic reticulum.

(D) Outline of the SILAC-based quantitative MS workflow employed in subtractive proteomics experiments. Equal amounts of crude and pure mitochondrial fractions prepared from differentially labeled HEK293T cells as depicted in (B) were mixed and proteins were digested with different proteases. Peptide mixtures were separated by high pH reversed-phase chromatography (pH 10 RP) into 32 fractions (fr.) per replicate (n = 4) and analyzed by LC-MS. Chymo., chymotrypsin.

(E) Reproducibility of protein quantification between four biological replicates. Numbers indicate the Pearson correlation between $\log_2$-transformed protein abundance ratios (pM/cM) of the indicated replicates, revealing a high correlation of SILAC ratios between replicates.

(F) Overlap of proteins quantified in four biological replicates. A total of 8,713 unique proteins were quantified with an average of 7,896 (± 154) per replicate.

(G) Enrichment (log_2 ratios > 0) or depletion (log_2 ratios < 0) of mitochondrial (red) and other proteins (blue) from different subcellular compartments in pure compared to crude mitochondrial fractions. Protein ratios were filtered for ≥ 3 valid values (n=4). MitoCoP refers to the high-confidence human mitochondrial proteome defined in this study (see Table S1). The 'Mito test set' contains known mitochondrial proteins according to the integrated mitochondrial protein index (IMPI). Non-mitochondrial subsets are based on GO-annotations excluding MitoCoP components. cM/pM, crude/pure mitochondria.

(H) GO-CC term overrepresentation analysis of classes defined in Figure 1B. Shown are selected overrepresented terms and the number of proteins assigned to a given term that are present in the respective class and, in parentheses, in the entire subtractive proteomics dataset. Numbers of proteins indicated in G and following subfigures include isoforms. RC, respiratory chain; PM, plasma membrane; ER, endoplasmic reticulum.

(I) Ratio-intensity plot as shown in Figure 1B illustrating the distribution of proteins of the mitochondrial (left) and non-mitochondrial assessment ('test') set (middle) as provided by MitoMiner/IMPI. The bar chart (right) shows the distribution of proteins of the subtractive proteomics dataset ('classification') and the mitochondrial and non-mitochondrial test sets across the classes defined in Figure 1B. Numbers indicate numbers of proteins present in the respective class.

(J, K) Same as in (I) highlighting proteins of the mitochondrial test set with amino-terminal presequence as predicted by MitoFates (Fukasawa et al., 2015) (J) and mitochondrial proteins of distinct submitochondrial localizations (K). The bar chart in (K) illustrates the distribution of proteins with indicated submitochondrial localization (based on GO-CC) across the classes defined in Figure 1B. Numbers indicate numbers of proteins present in the respective class. OM/IM, outer/inner mitochondrial membrane; IMS, intermembrane space.

(L) Distribution of components of the TIM22 complex, OXPHOS complexes, proteins involved in mitochondrial fission and fusion, components of mitochondrial ribosomes, mitochondrial proteins with dual localization and aminoacyl-tRNA synthetases in the subtractive proteomics dataset. Class 1 contains dual-localized proteins of predominantly mitochondrial localization, *e.g.* CHCHD2, COQ7 and HMGCL. Dually/multiply localized

proteins of minor abundance in mitochondria occur in class 2 such as CYB5R3, PARK7 and GARS. The log_2 SILAC ratio of >0 determined for these proteins points to the prevalence of the mitochondrial over the non-mitochondrial portion in the mitochondrial fractions analyzed in this experiment. loc., localized; AA, aminoacyl.

(H - L) Colors indicate classes as introduced in Figure 1B.

See also Tables S1 and S2.

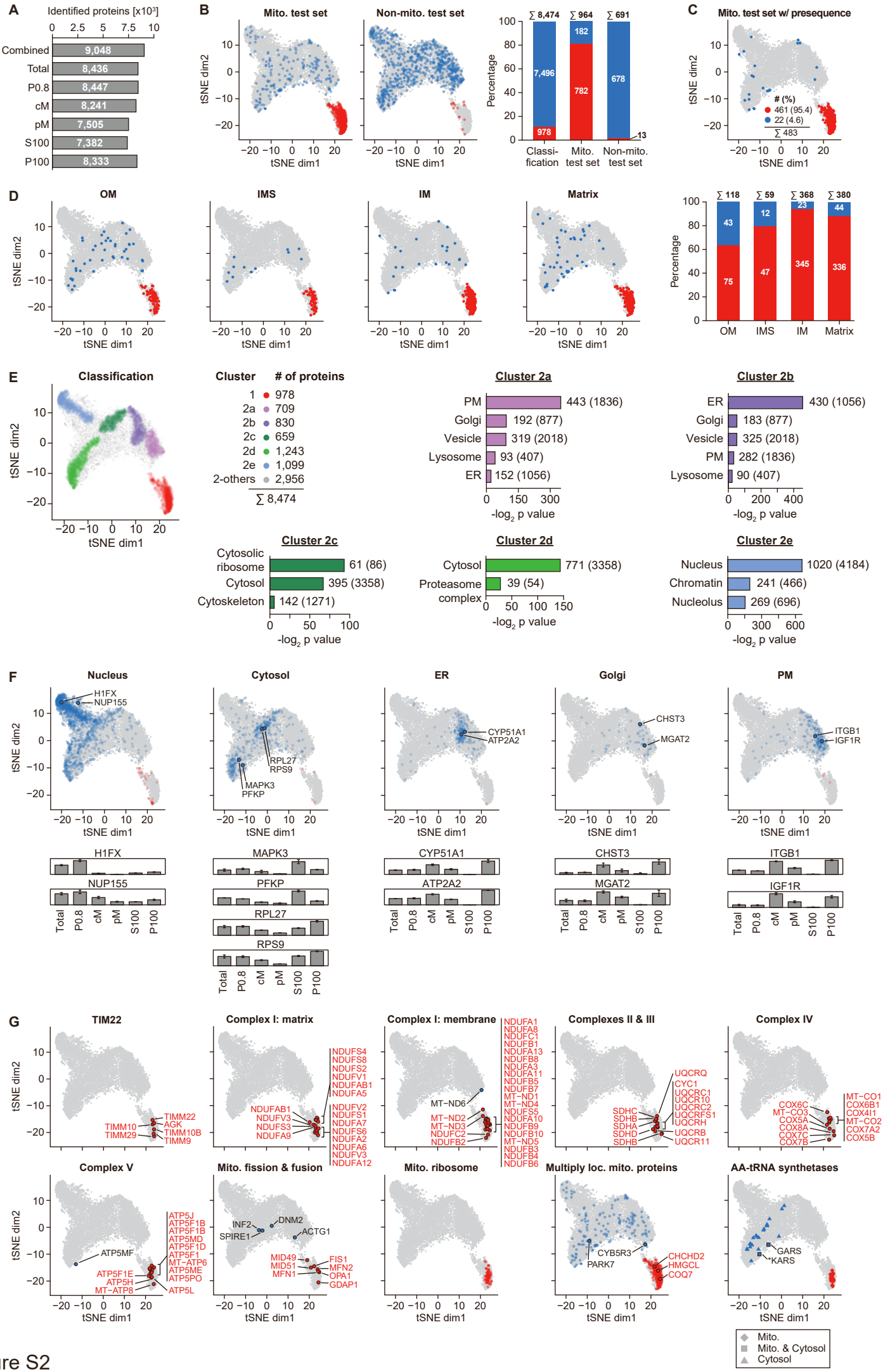
A
Identified proteins [x10³]
0 2.5 5 7.5 10
Combined 9,048
Total 8,436
P0.8 8,447
cM 8,241
pM 7,505
S100 7,382
P100 8,333
B
Mito. test set
Non-mito. test set
tSNE dim1
tSNE dim2
∑ 8,474
∑ 964
∑ 691
Percentage
7,496
978
182
782
678
13
Classi-fication
Mito. test set
Non-mito. test set
C
Mito. test set w/ presequence
(%)
461 (95.4)
22 (4.6)
∑ 483
D
OM
IMS
IM
Matrix
∑ 118
∑ 59
∑ 368
∑ 380
43
75
12
47
23
345
44
336
OM
IMS
IM
Matrix
E
Classification
Cluster # of proteins
1 978
2a 709
2b 830
2c 659
2d 1,243
2e 1,099
2-others 2,956
∑ 8,474
Cluster 2a
PM 443 (1836)
Golgi 192 (877)
Vesicle 319 (2018)
Lysosome 93 (407)
ER 152 (1056)
-log2 p value
Cluster 2b
ER 430 (1056)
Golgi 183 (877)
Vesicle 325 (2018)
PM 282 (1836)
Lysosome 90 (407)
Cluster 2c
Cytosolic ribosome 61 (86)
Cytosol 395 (3358)
Cytoskeleton 142 (1271)
Cluster 2d
Cytosol 771 (3358)
Proteasome complex 39 (54)
Cluster 2e
Nucleus 1020 (4184)
Chromatin 241 (466)
Nucleolus 269 (696)
F
Nucleus
H1FX
NUP155
Cytosol
MAPK3
RPL27
RPS9
PFKP
ER
CYP51A1
ATP2A2
Golgi
CHST3
MGAT2
PM
ITGB1
IGF1R
Total P0.8 cM pM S100 P100
G
TIM22
TIMM22
AGK
TIMM10
TIMM10B
TIMM29
TIMM9
Complex I: matrix
NDUFS4 NDUFS8 NDUFS2 NDUFV1 NDUFAB1 NDUFA5
NDUFAB1 NDUFV3 NDUFS3 NDUFA9
NDUFV2 NDUFS1 NDUFA7 NDUFS6 NDUFA2 NDUFA6 NDUFV3 NDUFA12
Complex I: membrane
NDUFA1 NDUFA8 NDUFC1 NDUFB1 NDUFA13 NDUFB8 NDUFA3 NDUFA11 NDUFB5 NDUFB7 MT-ND1 MT-ND4 NDUFS5 NDUFA10 NDUFB9 NDUFB10 MT-ND5 NDUFB3 NDUFB4 NDUFB6
MT-ND6
MT-ND2
MT-ND3
NDUFC2
NDUFB2
Complexes II & III
UQCRQ CYC1 UQCRC1 UQCR10 UQCRC2 UQCRFS1 UQCRH UQCRB UQCR11
SDHC SDHB SDHA SDHD SDHB
Complex IV
MT-CO1 COX6B1 COX4I1 MT-CO2 COX7A2 COX5B
COX6C MT-CO3 COX5A COX8A COX7C COX7B
Complex V
ATP5J ATP5F1B ATP5F1B ATP5MD ATP5F1D ATP5F1 MT-ATP6 ATP5ME ATP5PO ATP5L
ATP5MF
ATP5F1E ATP5H MT-ATP8
Mito. fission & fusion
DNM2
INF2
SPIRE1
ACTG1
MID49 MID51 MFN1 FIS1 MFN2 OPA1 GDAP1
Mito. ribosome
Multiply loc. mito. proteins
CYB5R3
PARK7
CHCHD2 HMGCL COQ7
AA-tRNA synthetases
GARS
KARS
Mito.
Mito. & Cytosol
Cytosol

Figure S2

Figure S2. Analysis of Spatial Proteomics Data

Related to Figure 1

(A) Number of proteins including isoforms present in the entire spatial proteomics dataset ('Combined') and identified in individual subcellular fractions generated as described in Figure S1B and analyzed by quantitative MS in a label-free approach (n = 3). Only proteins with MS intensity in ≥ 2/3 replicates per fraction were counted. Total, cell lysate; P0.8, nuclear fraction; cM/pM, crude/pure mitochondria; S100/P100, cytosolic/microsomal fraction.

(B) tSNE plots as shown in Figure 1D depicting the distribution of proteins of the mitochondrial (left) and non-mitochondrial test set (middle). Numbers in the bar chart (right) indicate the number of proteins assigned to cluster 1 (red) and cluster 2 (blue). Only 13 proteins (1.9%) of the non-mitochondrial test set were assigned to class 1 illustrating the high potential of the spatial proteomics approach to discriminate between mitochondrial and non-mitochondrial proteins. dim, dimension.

(C, D, F, G) Same as in (B) highlighting proteins of the mitochondrial test set with amino-terminal presequence as predicted by MitoFates (Fukasawa et al., 2015) (C), proteins of distinct submitochondrial (D) and subcellular localizations (F), and proteins of the TIM22 complex, OXPHOS complexes, proteins involved in mitochondrial fission and fusion, components of mitochondrial ribosomes, mitochondrial proteins with dual localization and aminoacyl-tRNA synthetases (G). All OXPHOS subunits highlighted in (G) were assigned to cluster 1; exceptions are MT-ND6 and ATP5MF, which exhibit limited or unfavorable cleavage sites for tryptic peptide-based identification and quantification. The bar chart in (D) indicates the number of proteins of distinct submitochondrial localizations assigned to cluster 1 (red) and cluster 2 (blue). Abundance profiles of selected proteins of different subcellular compartments are shown in (F). The localization of these proteins in the respective tSNE plots is indicated. ER, endoplasmic reticulum; OM/IM; outer/inner mitochondrial membrane; IMS, intermembrane space; PM, plasma membrane; Total, cell lysate; P0.8, nuclear fraction; cM/pM, crude/pure mitochondria; S100/P100, cytosolic/microsomal fraction; loc., localized; AA, aminoacyl.

(E) Left, same as in Figure 1D highlighting proteins of cluster 1 (*i.e.* the mitochondrial cluster) and different subclusters, which were defined by density-based clustering of the non-mitochondrial cluster 2. Right, GO-CC term overrepresentation analysis of subclusters 2a-2e. Shown are selected overrepresented terms including the number of proteins assigned to a given term that are present in the respective class and, in parentheses, in the entire subtractive proteomics dataset. Numbers of proteins indicated in G and following subfigures include isoforms. PM, plasma membrane; ER, endoplasmic reticulum.

(B-D, F, G) Colors indicate clusters as depicted in Figure 1D.

(F) Error bars indicate SEM for n = 3 and the range for n = 2.

See also Tables S1 and S2.

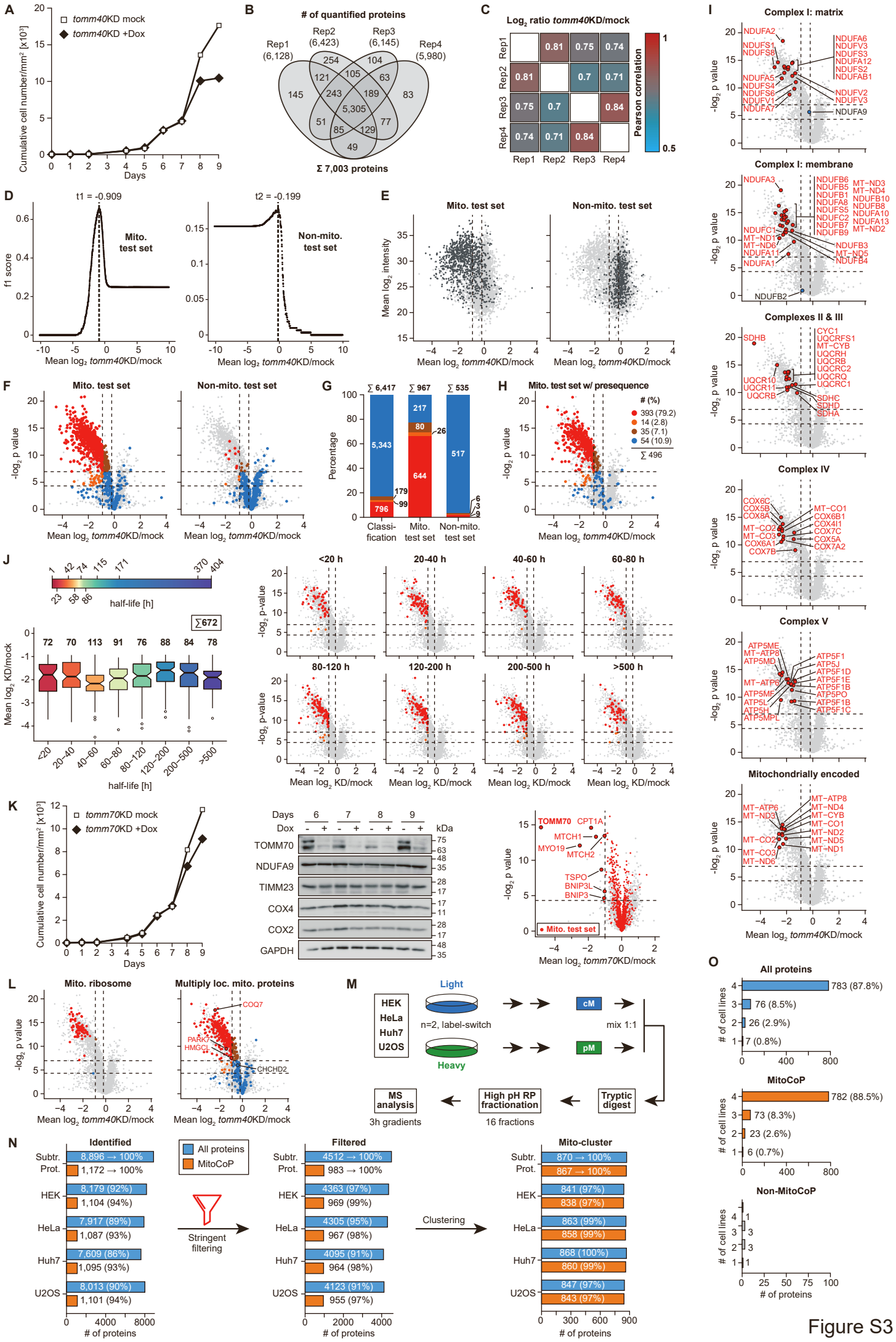
A
tomm40KD mock
tomm40KD +Dox
Cumulative cell number/mm² [x10³]
Days
B
# of quantified proteins
Rep1 (6,128)
Rep2 (6,423)
Rep3 (6,145)
Rep4 (5,980)
Σ 7,003 proteins
C
Log2 ratio tomm40KD/mock
Pearson correlation
D
t1 = -0.909
t2 = -0.199
Mito. test set
Non-mito. test set
f1 score
Mean log2 tomm40KD/mock
E
Mito. test set
Non-mito. test set
Mean log2 intensity
F
-log2 p value
G
Σ 6,417
Σ 967
Σ 535
Percentage
Classi-fication
H
Mito. test set w/ presequence
(%)
393 (79.2)
14 (2.8)
35 (7.1)
54 (10.9)
Σ 496
I
Complex I: matrix
Complex I: membrane
Complexes II & III
Complex IV
Complex V
Mitochondrially encoded
J
half-life [h]
Σ672
Mean log2 KD/mock
<20 h
20-40 h
40-60 h
60-80 h
80-120 h
120-200 h
200-500 h
>500 h
-log2 p-value
K
tomm70KD mock
tomm70KD +Dox
Days
Dox
kDa
TOMM70
NDUFA9
TIMM23
COX4
COX2
GAPDH
Mean log2 tomm70KD/mock
L
Mito. ribosome
Multiply loc. mito. proteins
M
HEK
HeLa
Huh7
U2OS
Light
Heavy
n=2, label-switch
cM
pM
mix 1:1
Tryptic digest
High pH RP fractionation
16 fractions
MS analysis
3h gradients
N
Identified
All proteins
MitoCoP
Stringent filtering
Filtered
Clustering
Mito-cluster
Subtr. Prot.
# of proteins
O
All proteins
MitoCoP
Non-MitoCoP
# of cell lines
# of proteins

Figure S3

Figure S3. Evaluation of Importomics and Multiple Cell Line Subtractive Proteomics Experiments

Related to Figure 2

(A) Representative growth curves of *tomm40*-shRNA cells treated with doxycycline (Dox) to induce *tomm40* shRNA or mock-treated with DMSO (control). Dox was replenished every 48 to 72 hours to continue shRNA induction. Cells were treated with Dox or DMSO for 9 days. KD, shRNA-mediated knockdown.

(B) Overlap of proteins including isoforms quantified in four biological replicates.

(C) Reproducibility of protein abundance ratios determined in four biological replicates. Values indicate the Pearson correlation between $\log_2$-transformed *tomm40*KD/mock ratios of replicates.

(D) Computation of significance thresholds for the definition of mitochondrial proteins based on the distribution of mitochondrial and non-mitochondrial test sets. F1 scores were determined for defined mean $\log_2$ *tomm40*KD/mock ratios using the mitochondrial (left) and non-mitochondrial test set (right). The ratio with the highest F1 score best separates the test set tested from the opposite test set. t1 and t2 represent the mean $\log_2$ ratios best separating the mitochondrial from the non-mitochondrial test set and *vice versa* in this dataset.

(E) Ratio-intensity plots illustrating the distribution of mitochondrial (left) and non-mitochondrial test proteins (right), each highlighted in dark grey.

(F) Volcano plots as depicted in Figure 2C showing the distribution of mitochondrial (left) and non-mitochondrial test proteins (right).

(G) Distribution of proteins of the TOMM40 importomics dataset ('Classification') and the mitochondrial and non-mitochondrial test set across the classes depicted in Figure 2C. The majority of mitochondrial test proteins (644) was present in class 1 and 106 proteins were grouped into classes 2 and 3. Only 9 proteins of the non-mitochondrial test set were found in class 1, while 517 non-mitochondrial proteins were members of class 4.

(H) Same as in (F) illustrating the distribution of proteins of the mitochondrial test set with amino-terminal presequence as predicted by MitoFates.

(I) Same as in (F) highlighting proteins of OXPHOS complexes.

(J) Assessment of a correlation between mitochondrial protein half-lives and protein abundance ratios determined in TOMM40 importomics experiments. The analysis is based on protein dynamics data generated in this study using HeLa cells (see Figures 7 and S7 and Table S6) for proteins of the high-confidence human mitochondrial proteome MitoCoP. Class I and class II proteins of the importomics dataset were grouped into distinct half-life bins as indicated. Boxplots show the distribution of KD/mock abundance ratios within a given bin. Numbers above boxplots indicate the number of proteins per bin. Volcano plots highlight the distribution of proteins with the indicated range of half-lives. Colors indicate classes as introduced in Figure 2C. KD, knockdown.

(K) Results of TOMM70 importomics experiments, performed as described in Figure 2B for TOMM40. Shown are representative growth curves of Dox- and mock-treated *tomm70*-shRNA cells treated as described in (A) for *tomm40*-shRNA cells (left); Western blot analysis of steady-state levels of TOMM70, selected other mitochondrial proteins and GAPDH (loading control) in whole cell lysates of Dox-treated (+) and untreated (-) *tomm70*-shRNA cells (middle); and the effect of *tomm70* knockdown on the mitochondrial proteome (proteins depicted were quantified in ≥ 3/4 replicates) (right).

(L) Same as in (F) highlighting components of mitochondrial ribosomes and mitochondrial proteins with dual localization. The data confirm a mitochondrial localization of the dual-localized proteins COQ7, PARK7, and HMGCL (see Figures S1L and S2F). loc., localized.

(M) Experimental setup of the multiple cell line subtractive proteomics approach. Cells of indicated cell lines were SILAC-labeled, followed by isolation of crude and pure mitochondrial fractions (n = 2). For each replicate, equal amounts of differentially labeled crude and pure mitochondrial fractions were mixed and proteolytically cleaved using trypsin. Tryptic peptides were fractionated and analyzed by LC-MS. cM/pM, crude/pure mitochondrial fraction.

(N) Determination of high-confidence mitochondrial proteins in multiple cell lines. Data acquired in the multiple cell line approach were analyzed together with the data of the large-scale multi-protease subtractive proteomics experiment of HEK293T cells ('Subtr. Prot.'; see Figures 1B, 1C, S1 and Table S2). Stringent filtering of proteins identified included (i) quantification in each replicate of the multiple cell line approach and

(ii) a sequence coverage of ≥ 10% in each replicate per dataset. To define sets of mitochondrial proteins ('Mito-cluster'), data were subjected to a density-based clustering approach as detailed in STAR Methods. 'All proteins' refers to MitoCoP and non-MitoCoP proteins.

(O) Conformity of Mito-Cluster components as defined in (N) across cell lines. ‘All proteins’ corresponds to MitoCoP and non-MitoCoP proteins.

(F - I, L) Colors indicate classes as introduced in Figure 2C.

See also Tables S1, S2, S3 and S6.

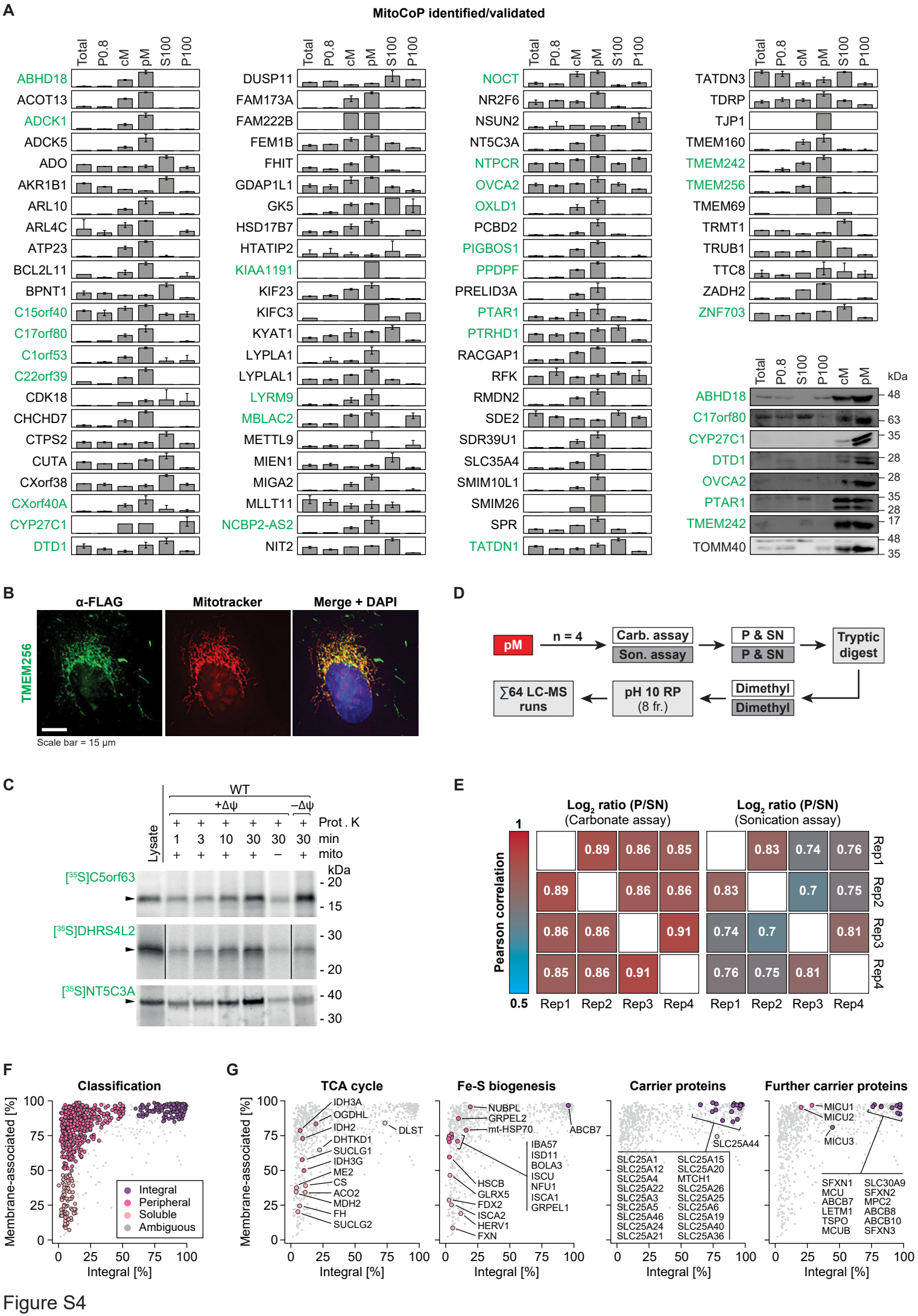
A
MitoCoP identified/validated
Total P0.8 cM pM S100 P100
ABHD18 ACOT13 ADCK1 ADCK5 ADO AKR1B1 ARL10 ARL4C ATP23 BCL2L11 BPNT1 C15orf40 C17orf80 C1orf53 C22orf39 CDK18 CHCHD7 CTPS2 CUTA CXorf38 CXorf40A CYP27C1 DTD1
DUSP11 FAM173A FAM222B FEM1B FHIT GDAP1L1 GK5 HSD17B7 HTATIP2 KIAA1191 KIF23 KIFC3 KYAT1 LYPLA1 LYPLAL1 LYRM9 MBLAC2 METTL9 MIEN1 MIGA2 MLLT11 NCBP2-AS2 NIT2
NOCT NR2F6 NSUN2 NT5C3A NTPCR OVCA2 OXLD1 PCBD2 PIGBOS1 PPDPF PRELID3A PTAR1 PTRHD1 RACGAP1 RFK RMDN2 SDE2 SDR39U1 SLC35A4 SMIM10L1 SMIM26 SPR TATDN1
TATDN3 TDRP TJP1 TMEM160 TMEM242 TMEM256 TMEM69 TRMT1 TRUB1 TTC8 ZADH2 ZNF703
Total P0.8 S100 P100 cM pM kDa
ABHD18 48
C17orf80 63
CYP27C1 35
DTD1 28
OVCA2 28
PTAR1 35 28
TMEM242 17
TOMM40 48 35
B
α-FLAG Mitotracker Merge + DAPI
TMEM256
Scale bar = 15 µm
C
WT
+Δψ −Δψ
Lysate
Prot . K + + + + + +
1 3 10 30 30 30 min
mito + + + + − +
kDa
[35S]C5orf63 - 20 - 15
[35S]DHRS4L2 - 30 - 20
[35S]NT5C3A - 40 - 30
D
pM n = 4 Carb. assay Son. assay P & SN P & SN Tryptic digest
Dimethyl Dimethyl pH 10 RP (8 fr.) ∑64 LC-MS runs
E
Log2 ratio (P/SN) (Carbonate assay)
Log2 ratio (P/SN) (Sonication assay)
Pearson correlation 1 0.5
0.89 0.86 0.85
0.89 0.86 0.86
0.86 0.86 0.91
0.85 0.86 0.91
0.83 0.74 0.76
0.83 0.7 0.75
0.74 0.7 0.81
0.76 0.75 0.81
Rep1 Rep2 Rep3 Rep4
F
Classification
Membrane-associated [%]
Integral [%]
Integral
Peripheral
Soluble
Ambiguous
G
TCA cycle
IDH3A OGDHL IDH2 DLST DHTKD1 SUCLG1 IDH3G ME2 CS ACO2 MDH2 FH SUCLG2
Fe-S biogenesis
NUBPL GRPEL2 mt-HSP70 ABCB7 IBA57 ISD11 BOLA3 ISCU NFU1 ISCA1 GRPEL1 HSCB GLRX5 FDX2 ISCA2 HERV1 FXN
Carrier proteins
SLC25A44
SLC25A1 SLC25A12 SLC25A4 SLC25A22 SLC25A3 SLC25A5 SLC25A46 SLC25A24 SLC25A21
SLC25A15 SLC25A20 MTCH1 SLC25A26 SLC25A25 SLC25A6 SLC25A19 SLC25A40 SLC25A36
Further carrier proteins
MICU1 MICU2 MICU3
SFXN1 MCU ABCB7 LETM1 TSPO MCUB
SLC30A9 SFXN2 MPC2 ABCB8 ABCB10 SFXN3

Figure S4

Figure S4. Analysis of Subcellular Localization and Membrane Association of MitoCoP Proteins

Related to Figure 3

(A) Protein correlation profiles for MitoCoP-A/B proteins established in spatial proteomics experiments. Shown are the mean of normalized MS intensities (top) and Western blot analyses (bottom) of the indicated fractions. Error bars indicate SEM for n = 3 and the range for n = 2. Total, cell lysate; P0.8, nuclear fraction; cM/pM, crude/pure mitochondria; S100/P100, cytosolic/microsomal fraction. Immunodecoration of cellular fractions by specific antibodies.

(B) Immunofluorescence microscopy of transiently transfected U2OS cells expressing $TMEM256_{FLAG}$. $TMEM256_{FLAG}$ was visualized by anti-FLAG antibody, the mitochondrial network by Mitotracker red staining and the nucleus by DAPI. Scale bar, 15 μm.

(C) *In organello* import assays of radiolabeled precursor proteins into HEK293T mitochondria in the presence or absence of a membrane potential ($\Delta\psi$) or mitochondria, followed by proteinase K (Prot. K) treatment. Lysate, *in vitro* synthesized radiolabeled precursor protein.

(D) Experimental strategy followed to assess the membrane association of MitoCoP proteins. Gradient-purified mitochondria were either treated with sodium carbonate pH 11.5 to separate integral membrane proteins from other proteins (*i.e.* carbonate assay) or sonified to separate integral and peripheral membrane proteins from soluble proteins (*i.e.* sonication assay). Proteins of soluble and insoluble fractions generated by centrifugation were digested with trypsin and the resulting peptides were labeled using stable isotope dimethyl labeling. Differentially labeled peptides from soluble and insoluble fractions of each replicate (n = 4) were mixed, fractionated by high pH reversed-phase chromatography (pH 10 RP) and analyzed by LC-MS. Carb., carbonate; Son., sonication; P, pellet/insoluble fraction; SN, supernatant/soluble fraction; fr., fractions.

(E) Reproducibility of protein quantification between four biological replicates of carbonate (left) and sonication assays (right). Numbers indicate the Pearson correlation between log_2-transformed P/SN protein abundance ratios of the indicated replicates.

(F) Membrane association of MitoCoP proteins. Carbonate extraction and sonication of gradient-purified mitochondria as described in (C) results in the separation of the mitochondrial proteome into integral membrane proteins, peripheral membrane proteins and soluble proteins as reflected in this scatterplot showing the classification. Classes were defined by statistical data analysis based on values for '% associated' and '% integral'. 'Associated' and 'Integral' denote data of sonication and carbonate extraction experiments, respectively.

(G) Scatterplots as shown in (F) highlighting mitochondrial carrier proteins, proteins of the tricarboxylic acid cycle and proteins involved in iron-sulfur cluster biogenesis. Colors indicate classes as shown in (D). TCA, tricarboxylic acid; Fe-S, iron-sulfur.

See also Tables S1 and S3.

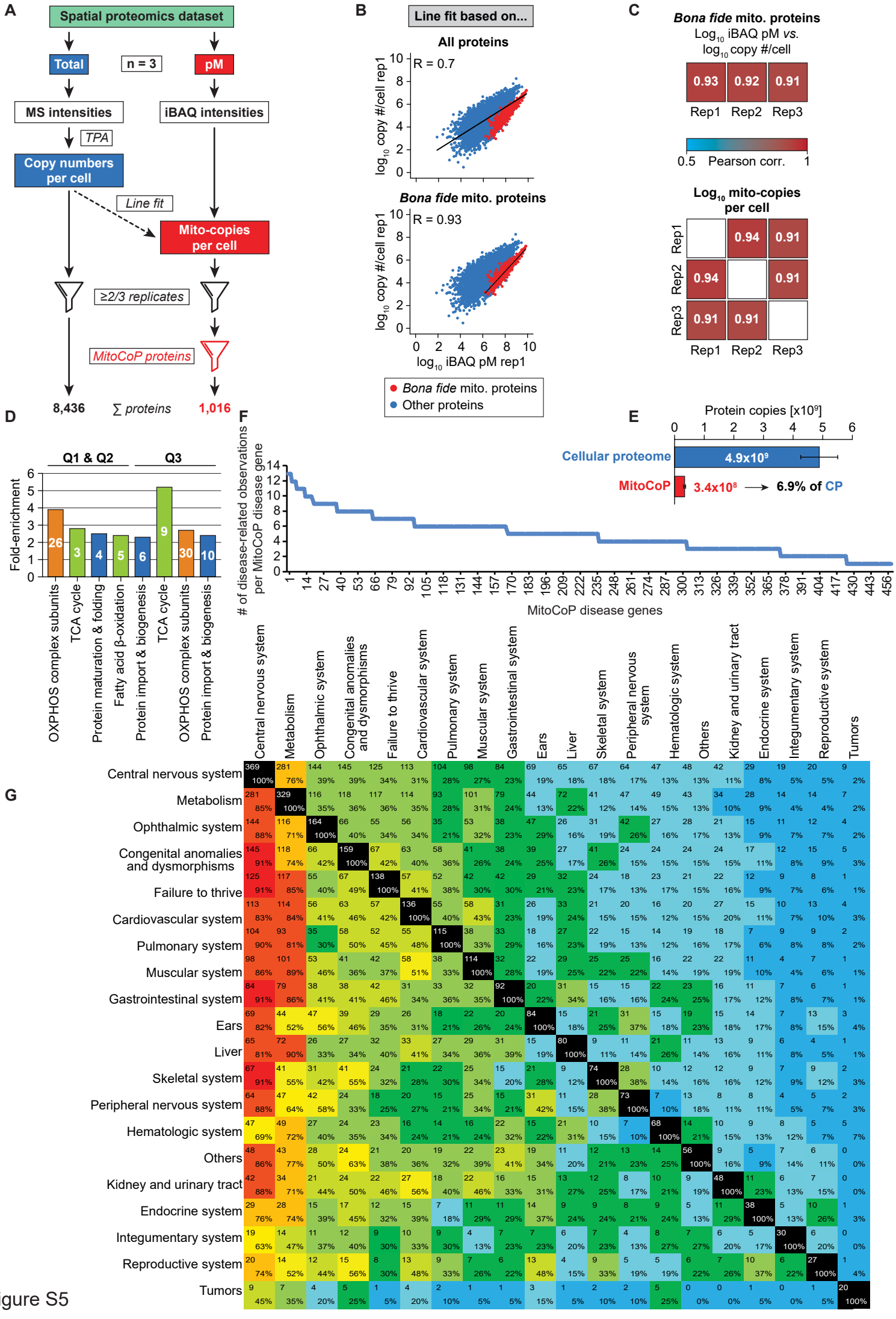
A
Spatial proteomics dataset
Total
n = 3
pM
MS intensities
iBAQ intensities
TPA
Copy numbers per cell
Line fit
Mito-copies per cell
≥2/3 replicates
MitoCoP proteins
8,436
∑ proteins
1,016
B
Line fit based on...
All proteins
R = 0.7
log10 copy #/cell rep1
Bona fide mito. proteins
R = 0.93
log10 iBAQ pM rep1
Bona fide mito. proteins
Other proteins
C
Bona fide mito. proteins
Log10 iBAQ pM vs. log10 copy #/cell
0.93
0.92
0.91
Rep1
Rep2
Rep3
0.5 Pearson corr. 1
Log10 mito-copies per cell
0.94
0.91
D
Q1 & Q2
Q3
Fold-enrichment
OXPHOS complex subunits
TCA cycle
Protein maturation & folding
Fatty acid β-oxidation
Protein import & biogenesis
TCA cycle
OXPHOS complex subunits
Protein import & biogenesis
E
Protein copies [x10⁹]
Cellular proteome
4.9x10⁹
MitoCoP
3.4x10⁸
6.9% of CP
F
# of disease-related observations per MitoCoP disease gene
MitoCoP disease genes
G
Central nervous system
Metabolism
Ophthalmic system
Congenital anomalies and dysmorphisms
Failure to thrive
Cardiovascular system
Pulmonary system
Muscular system
Gastrointestinal system
Ears
Liver
Skeletal system
Peripheral nervous system
Hematologic system
Others
Kidney and urinary tract
Endocrine system
Integumentary system
Reproductive system
Tumors

Figure S5

Figure S5. Analysis of MitoCoP Protein Copy Numbers and Functional Disease Gene Classification

Related to Figures 4 and 5

(A) Outline of the strategy for the estimation of cellular protein copy numbers and mito-copy numbers per cell using the data acquired in the spatial proteomics approach. Overall cellular protein copy numbers were calculated using the MS intensities determined for whole cell lysates (Total) following the total protein approach (TPA) (Wiśniewski et al., 2012). To determine mito-copy numbers per cell, iBAQ intensities determined in gradient-purified mitochondrial fractions (pM) were used. Mito-copy numbers per cell were inferred from copy numbers per cell as shown in (B). iBAQ, intensity-based absolute quantification.

(B) Correlation between $\log_{10}$ copy numbers per cell and $\log_{10}$ iBAQ values obtained in the pM fraction determined for all proteins (left) and for *bona fide* mitochondrial proteins, defined as class 1 proteins of the spatial proteomics dataset after filtering (right). Shown are the data for replicate 1. R, Pearson correlation coefficient; rep, replicate.

(C) Reproducibility of the line-fit determined based on *bona fide* mitochondrial proteins (left) and of mito-copy numbers per cell (right) of all three biological replicates. corr., correlation; pM, pure mitochondria; rep, replicate.

(D) Enrichment of functional classes as defined in Figure 4A in the indicated quartiles (depicted in Figure 4B) with regard to the overall dataset, which comprises 1,016 mitochondrial protein groups for which mito-copy numbers per cell were calculated. Numbers indicate how many times a functional term occurs in the indicated quartile. Only terms with a fold enrichment of ≥ 2 are shown. Coloring refers to colors introduced in Figure 4A. TCA, tricarboxylic acid; Q, quartile.

(E) Summed protein copy numbers determined for the cellular proteome (CP, 8,436 proteins; including proteins of the MitoCoP) and for the mitochondrial proteome (MitoCoP, 1,016 proteins). Error bars indicate SEM.

(F) Detailed functional classification of the MitoCoP disease genes related to Figure 5B.

(G) Number of disease-related observations per MitoCoP gene.

See also Tables S1 and S4.

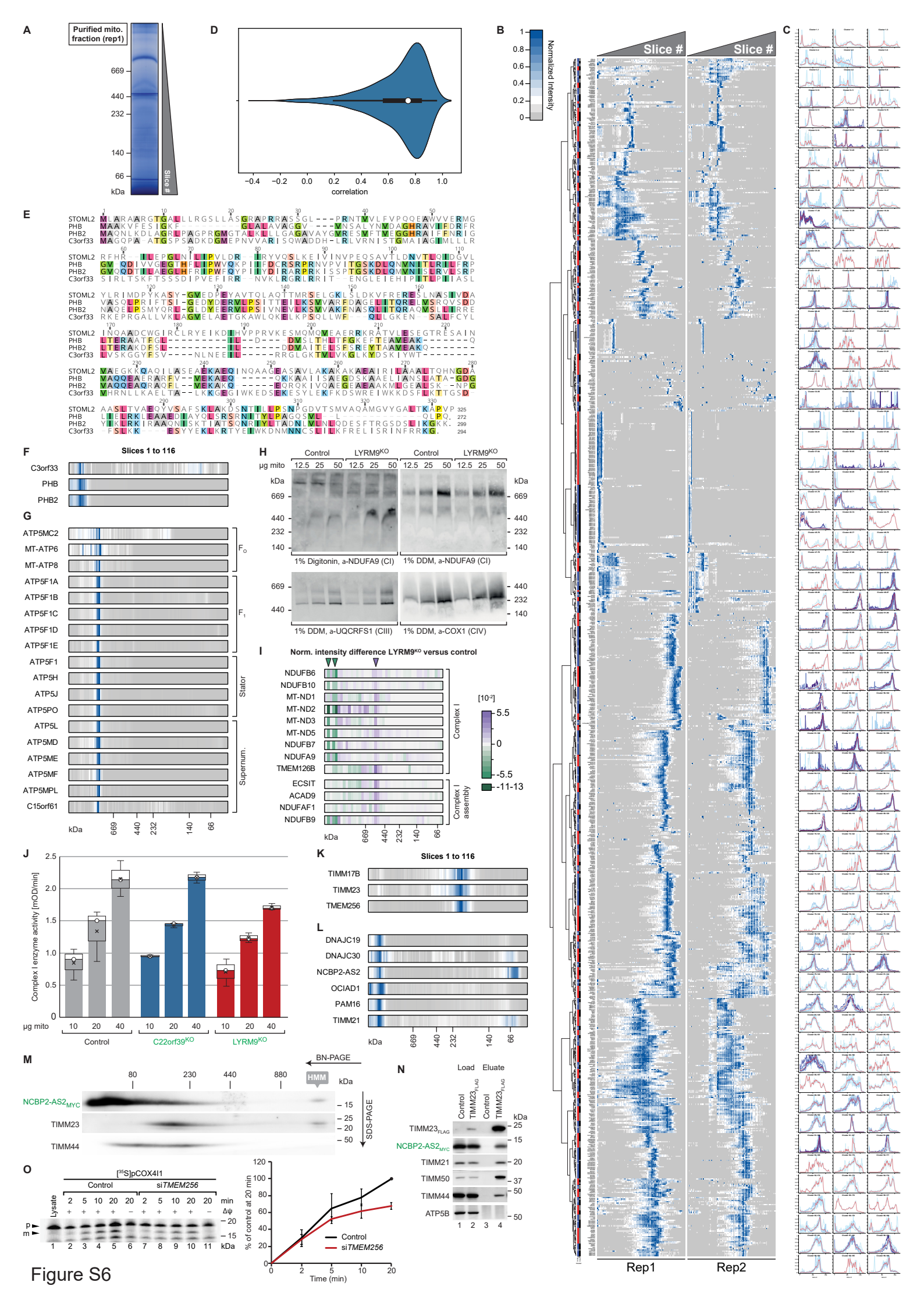
A
Purified mito. fraction (rep1)
669
440
232
140
66
kDa
Slice #
D
−0.4 −0.2 0.0 0.2 0.4 0.6 0.8 1.0
correlation
B
Normalized Intensity
1
0.8
0.6
0.4
0.2
0
Slice #
Slice #
C
E
STOML2
PHB
PHB2
C3orf33
F
Slices 1 to 116
C3orf33
PHB
PHB2
G
ATP5MC2
MT-ATP6
MT-ATP8
F0
ATP5F1A
ATP5F1B
ATP5F1C
ATP5F1D
ATP5F1E
F1
ATP5F1
ATP5H
ATP5J
ATP5PO
Stator
ATP5L
ATP5MD
ATP5ME
ATP5MF
ATP5MPL
C15orf61
Supernum.
kDa
669
440
232
140
66
H
Control
LYRM9KO
Control
LYRM9KO
µg mito
12.5 25 50 12.5 25 50 12.5 25 50 12.5 25 50
kDa
669
440
232
140
1% Digitonin, a-NDUFA9 (CI)
1% DDM, a-NDUFA9 (CI)
1% DDM, a-UQCRFS1 (CIII)
1% DDM, a-COX1 (CIV)
I
Norm. intensity difference LYRM9KO versus control
NDUFB6
NDUFB10
MT-ND1
MT-ND2
MT-ND3
MT-ND5
NDUFB7
NDUFA9
TMEM126B
ECSIT
ACAD9
NDUFAF1
NDUFB9
Complex I
Complex I assembly
[10-2]
5.5
0
-5.5
-11-13
kDa
669
440
232
140
66
J
Complex I enzyme activity [mOD/min]
2.5
2.0
1.5
1.0
0.5
0
µg mito
10 20 40 10 20 40 10 20 40
Control
C22orf39KO
LYRM9KO
K
Slices 1 to 116
TIMM17B
TIMM23
TMEM256
L
DNAJC19
DNAJC30
NCBP2-AS2
OCIAD1
PAM16
TIMM21
kDa
669
440
232
140
66
M
BN-PAGE
80
230
440
880
HMM
kDa
NCBP2-AS2MYC
TIMM23
TIMM44
15
25
20
50
SDS-PAGE
N
Load
Eluate
Control
TIMM23FLAG
Control
TIMM23FLAG
kDa
TIMM23FLAG
NCBP2-AS2MYC
TIMM21
TIMM50
TIMM44
ATP5B
25
15
20
37
50
50
1 2 3 4
O
[35S]pCOX4I1
Control
siTMEM256
Lysate
2 5 10 20 20 2 5 10 20 20 min
+ + + + − + + + + − Δψ
p
m
20
15
1 2 3 4 5 6 7 8 9 10 11 kDa
% of control at 20 min
120
100
80
60
40
20
0
0 2 5 10 20
Time (min)
Control
siTMEM256
Rep1
Rep2

Figure S6

Figure S6. Complexome Profiling of Human MitoCoP Proteins

Related to Figure 6

(A) Purified mitochondria isolated from human HEK293T cells were lysed with digitonin and separated by blue native (BN) PAGE (n = 2). Image shows the BN gel of replicate (rep) 1.

(B) Complexome profiles of purified mitochondria. Complexome #1 (Rep1, left) is based on LC-MS analysis of 116 slices cut from the whole BN gel shown in (A) and hierarchical clustering. For complexome #2 (Rep2, right), the BN gel was cut into 114 slices.

(C) Cluster profiles from complexome #1 (B), the rolling mean over three relative intensity values is depicted in red.

(D) Correlation of protein migration profiles between the replicates of the complexome analysis shown in (B). Comparison of 896 protein profiles resulted in a median correlation of 0.75.

(E) Sequence alignment highlighting the homology of C3orf33 with the prohibitin-stomatin protein family.

(F, G) Comparison of BN-profiles from prohibitins and C3orf33 (F) and subunits of the ATP synthase and C15orf61 (G).

(H) Immunodecorations of mitochondrial protein complexes resolved by BN-PAGE after lysis with digitonin or n-Dodecyl-β-D-maltoside (DDM).

(I) Label-free comparative complexome profiling of mitochondrial fractions isolated from HEK293T LYRM9^{KO} and control cells. Mitochondria were lysed with digitonin and separated by BN-PAGE. Lanes of both samples were cut into 45 slices and analyzed via LC-MS. Shown is the difference between normalized intensities of LYRM9^{KO} and control cells for the indicated proteins. Intensities were normalized by dividing each intensity by the sum of the respective row's intensity. Thus, the sum over all slices for an individual protein is identical for LYRM9^{KO} and control cells and set to 100% for the following examples. For the depicted proteins, the maximum intensity for an individual slice is 17.2% for control and 13.9% for LYRM9^{KO} cells. *E.g.* the 12.2% intensity difference in slice 2 for mt-ND2 between control (17.2%) and LYRM9^{KO} (5.0%) cells indicates a reduction to less than 30% mt-ND2 containing supercomplex. Conversely, the 5.0% difference in slice 20

for TMEM126B between LYRM9^{KO} (9.0%) and control cells (4.0%) indicates a ~225% increase of TMEM126B containing complex I assembly intermediate.

(J) Complex I activity of isolated C22orf39^{KO}, LYRM9^{KO} and control mitochondria. Box plot (n = 3) with whiskers extending to minimum and maximum.

(K, L) Comparison of BN-profiles from TIM23 core subunits and TMEM256 (K) and of proteins clustered to the TIM23 accessory subunits PAM and TIMM21 (L).

(M) NCBP2-AS2$_{MYC}$ was transfected transiently into HEK293 T cells. Isolated mitochondria were lysed with digitonin, protein complexes were separated by BN-PAGE, followed by a 2nd dimension SDS-PAGE and Western blot analysis. HMM, high molecular mass.

(N) NCBP2-AS2$_{Myc}$ was transiently transfected in TIM23$_{FLAG}$ expressing cells. FLAG-immunoprecipitations eluates of whole cells lysed with digitonin were analyzed by SDS-PAGE. (Load = 2%, Eluate = 100%).

(O) HEK293T cells were depleted for TMEM256 using siRNA molecules. After mitochondrial isolation, the [^{35}S]COX4I precursor was imported for indicated time points, followed by proteinase K treatment. Samples were analyzed by SDS-PAGE and digital autoradiography. Error bars, SEM (n = 3). Lysate, synthesized precursor; NT, not treated; p, precursor; m, mature.

See also Table S5.

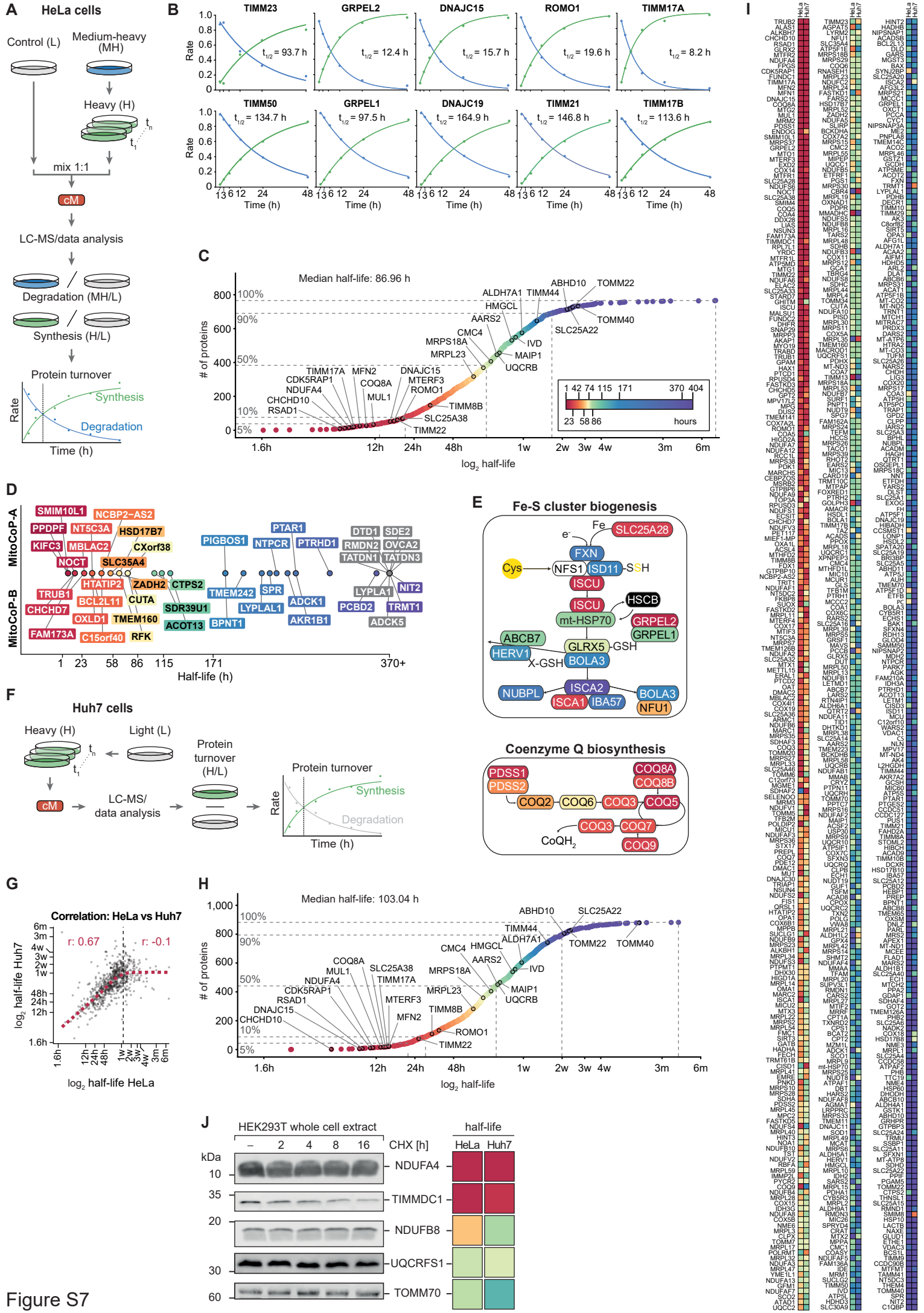
A
HeLa cells
Control (L)
Medium-heavy (MH)
Heavy (H)
mix 1:1
cM
LC-MS/data analysis
Degradation (MH/L)
Synthesis (H/L)
Protein turnover
Synthesis
Degradation
Rate
Time (h)
B
TIMM23 t1/2 = 93.7 h
GRPEL2 t1/2 = 12.4 h
DNAJC15 t1/2 = 15.7 h
ROMO1 t1/2 = 19.6 h
TIMM17A t1/2 = 8.2 h
TIMM50 t1/2 = 134.7 h
GRPEL1 t1/2 = 97.5 h
DNAJC19 t1/2 = 164.9 h
TIMM21 t1/2 = 146.8 h
TIMM17B t1/2 = 113.6 h
Rate
Time (h)
C
Median half-life: 86.96 h
# of proteins
log2 half-life
hours
D
MitoCoP-A
MitoCoP-B
Half-life (h)
E
Fe-S cluster biogenesis
Coenzyme Q biosynthesis
F
Huh7 cells
Heavy (H)
Light (L)
cM
LC-MS/ data analysis
Protein turnover (H/L)
Protein turnover
Synthesis
Degradation
Rate
Time (h)
G
Correlation: HeLa vs Huh7
r: 0.67
r: -0.1
log2 half-life Huh7
log2 half-life HeLa
H
Median half-life: 103.04 h
# of proteins
log2 half-life
I
HeLa
Huh7
J
HEK293T whole cell extract
CHX [h]
half-life
HeLa
Huh7
kDa
NDUFA4
TIMMDC1
NDUFB8
UQCRFS1
TOMM70

Figure S7

Figure S7. Evaluation of Pulsed SILAC Data for Human MitoCoP Half-life Analysis Related to Figure 7

(A) Outline of the pulsed SILAC (pSILAC) approach applied to determine half-lives of MitoCoP proteins, shown for replicate 1. HeLa cells were grown in cell culture medium containing light (L, control) or medium-heavy (MH) amino acids. 48 h, 24 h, 12 h, 6 h, 3 h, and 1 h prior to harvesting, MH-labeled cells were shifted to medium containing heavy (H) amino acids. Control cells were maintained in light medium for the entire duration of the experiment. Equal numbers of differentially SILAC-labeled pulsed and control cells were mixed, crude mitochondrial fractions were prepared and proteins were analyzed by LC-MS following digestion with trypsin and peptide fractionation using high pH reversed-phase chromatography. The calculation of protein synthesis and degradation rates was based on MH/L and H/L peptide/protein abundance ratios, respectively (this applies to replicate 1; for replicates 2 and 3, the respective label-switch has to be taken into account).

(B) Protein degradation (blue) and synthesis (green), exemplarily shown for selected proteins of the presequence translocase import motor system TIM23-PAM with fast and slow turnover rates determined by fitting exponential curves (lines) prior to correcting the data for cell doubling time. $T_{1/2}$, corresponding protein half-life.

(C) Range of half-lives determined for human MitoCoP constituents from HeLa cells.

(D, E) Half-lives determined for MitoCoP-A/B proteins (D) and proteins involved in iron-sulfur cluster biogenesis (E, top) and coenzyme Q biosynthesis (E, bottom).

(F) Outline of the pSILAC approach applied to determine half-lives of MitoCoP proteins from Huh7 cells. Following growth in cell culture medium containing light (L) amino acids, Huh7 cells were shifted to heavy (H) amino acids for 0 h, 5 h, 10 h, 24 h or 48 h prior to harvesting. Isolation of crude mitochondrial fractions was followed by peptide fractionation using high pH reversed-phase chromatography and LC-MS analysis (n = 2 for each time point). The calculation of protein synthesis and degradation rates was based on H/L peptide ratios. cM, crude mitochondrial fraction.

(G) Correlation of half-life values determined from pSILAC analysis of Huh7 and HeLa cells for proteins with half-lives in the range of 2 hours to 7.5 days (indicated via dashed black line) and above 7.5 days. r, Pearson correlation.

(H) Range of half-lives determined for human MitoCoP constituents from Huh7 cells.

(I) Comparison of half-lives of MitoCoP proteins determined from HeLa and Huh7 cells. Shown are 748 proteins for which protein half-lives were calculated from both cell lines.

(J) Western blotting of cell extracts grown for the indicated times in the presence of cycloheximide (CHX) compared to determined protein half-lives.

(D, E, H, J, I) color code as in (C).

See also Table S6.