Figure S1, related to Figure 1

Figure S1, related to Figure 1. Mitochondrial localization of baits used in this study.
Epifluorescence microscopy images showing the mitochondrial localization of all FLAG-tagged baits used in this study. Merged images show the mitochondria-localized GFP signal in green (Mito), the FLAG-tagged protein in red (FLAG) and the nucleus in blue (DNA). Regions of co-localization appear yellow.

Figure S2, related to Figure 2

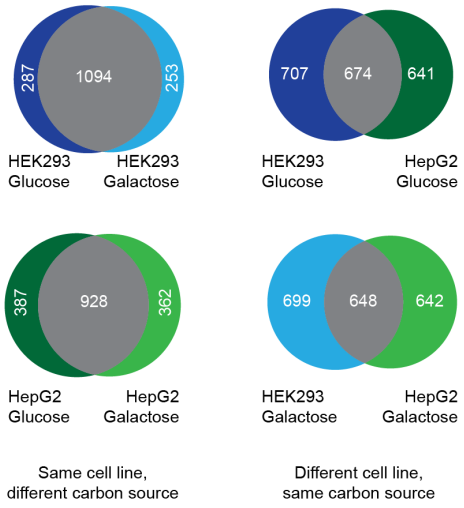


Figure S2, related to Figure 2. Effects of conditions and replicates on AE-MS analyses. Venn diagrams showing the interactions identified when only subsets of our data are used. Dark blue represents interactions identified in HEK293 cells grown in glucose, light blue is HEK293 in galactose, dark green is HepG2 in glucose, and light green is HepG2 in galactose. Grey represents the overlap in interactions between two groups.

Figure S3, related to Figure 3

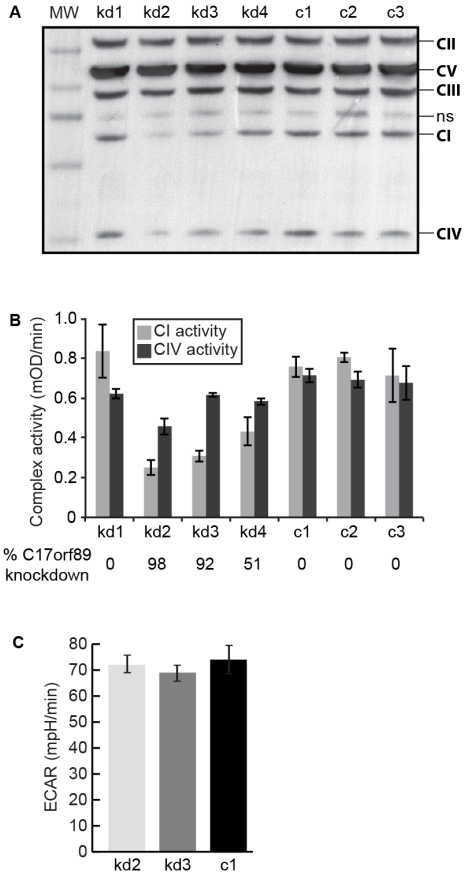
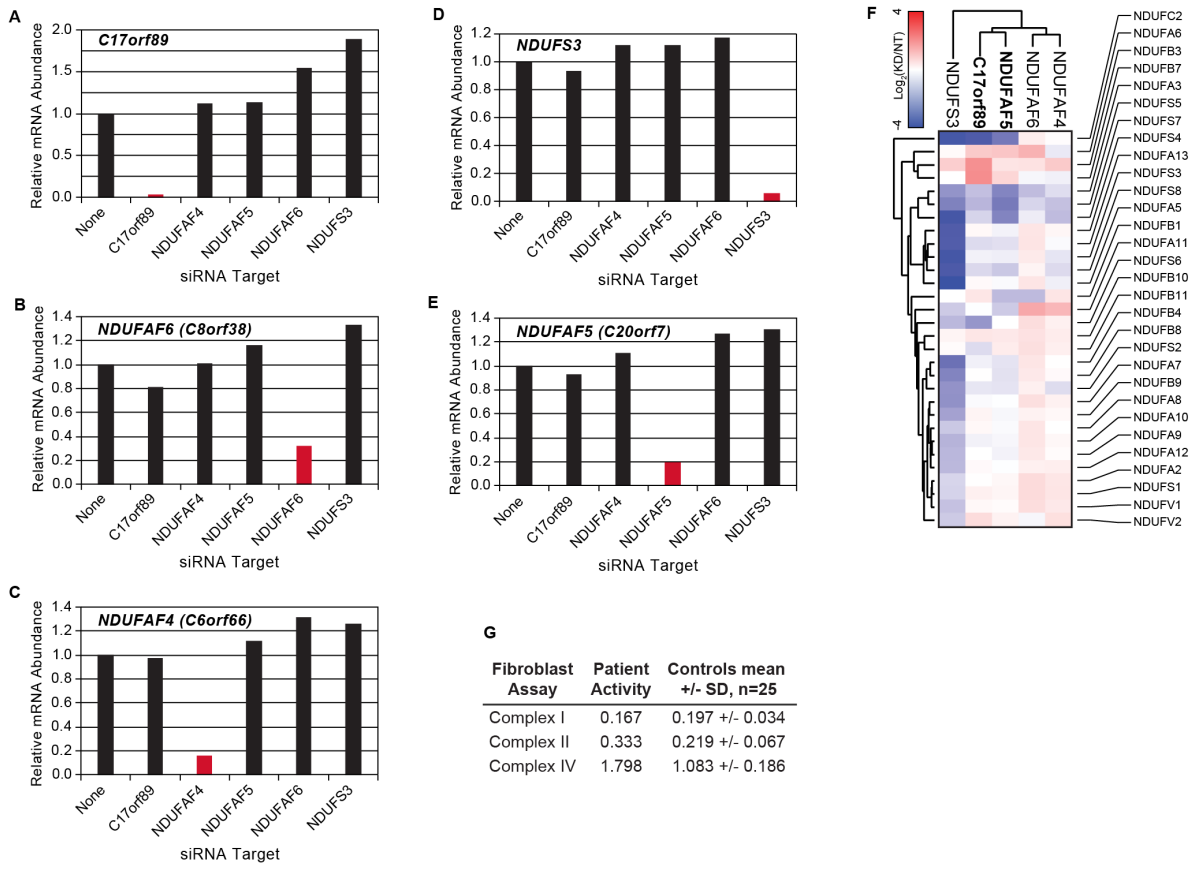


Figure S3, related to Figure 3. Functional assays implicating C17orf89 in CI assembly.
(A) Abundance of core OxPhos complex subunits in C17orf89 knock down (kd) and control (c) HEK293 cell lines (ns, non-specific band). Each gel lane is aligned to the activity data in (B).
(B) Complex I (CI) and complex IV (CIV) assays for C17orf89 knockdown (kd) and control (c) cells as described in Figure 6A of the main text. Error bars indicate ± standard deviation.
(C) Extracellular acidification rate (ECAR) measurements of C17orf89 knockdown (kd) and control (c) cells as described in Figure 3D of the main text. Error bars indicate ± standard error of the mean.

Figure S4, related to Figure 4



G

Fibroblast Assay	Patient Activity	Controls mean +/- SD, n=25
Complex I	0.167	0.197 +/- 0.034
Complex II	0.333	0.219 +/- 0.067
Complex IV	1.798	1.083 +/- 0.186

Figure S4, related to Figure 4. Knockdown of C17orf89 causes a tissue-specific CI defect that is similar to the loss of NDUFAF5.

(A-E) Relative mRNA abundance of the indicated siRNA targets across siRNA transfections as assessed by real-time qPCR. (F) Clustered heat map of CI protein subunit levels in CI or CIAF KD cells detected via MS. (G) Respiratory chain complex activity for fibroblasts from patient described in Figure 4.

Figure S5, related to Figure 5

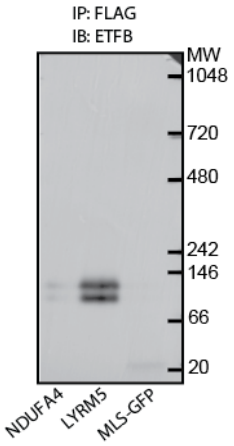
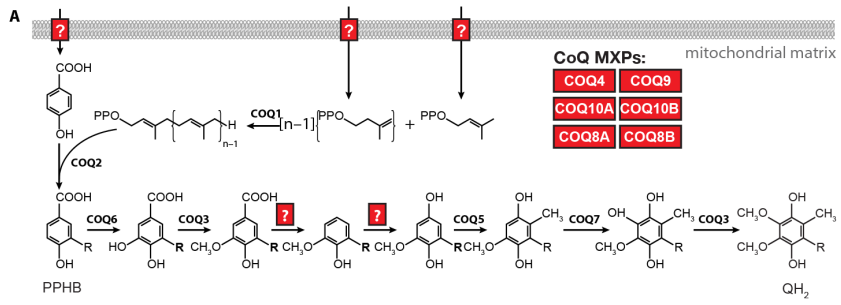


Figure S5, related to Figure 5. ETFB is also present in a complex immunoprecipitated by LYRM5/ETFBP.

The same sample IP eluate shown in Figure 5C was loaded onto a separate native gel and immunoblotted for ETFB.

Figure S6, related to Figure 7



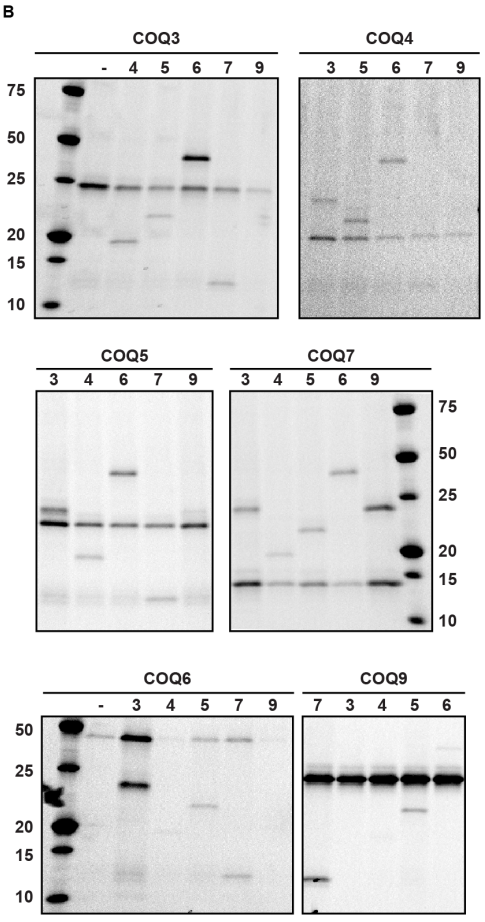


Figure S6, related to Figure 7. Analysis of binary interactions between CoQ-related proteins.

(A) Overview of the coenzyme Q biosynthesis pathway. Steps of the pathway that are not yet associated with a mitochondrial protein are indicated by a question mark (?). MXPs related to the CoQ pathway are boxed in red.

(B) Analyses of COQ protein pairwise interactions using wheat germ cell-free protein translation and purification. The protein listed above each gel was included in each lane. Numbers above each lane correspond to another COQ protein that was co-translated and co-purified with the protein listed above each gel (e.g., “4” indicates “COQ4”; “-“ indicates that no other protein was included). Numbers to the sides of gels are molecular weights in kilodaltons.

Figure S7, related to figures 3, 5, and 7

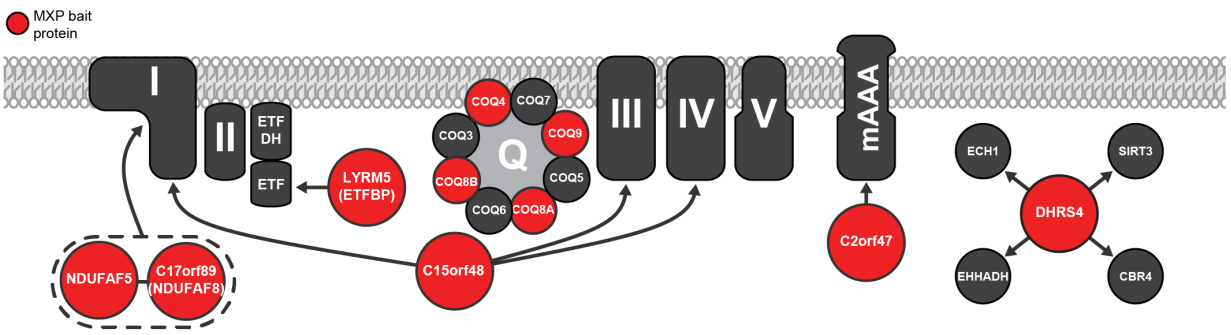


Figure S7, related to figures 3, 5, and 7. Overview of select interactions identified within our data. Red circles indicate MXPs used as bait proteins in this study. Arrows indicate interactions of MXPs with the corresponding mitochondrial proteins or complexes. Complex Q (Q), which involves many interactions identified in this study, is depicted in a simplified form.

SUPPLEMENTAL TABLE LEGENDS

Table S1. MitoCarta+ list, related to Figure 1A and 1B.
Table of our MitoCarta+ list of 1,166 human proteins with validated mitochondrial localization including the 50 MXPs and 27 mitochondrial proteins of known function used as baits in this study.

Table S2. Control PPI list, related to Figure 2A.
Table of literature-established protein-protein interactions (PPIs) involving our positive-control baits.

Table S3. PPIs established in this study, related to Figure 2.
Table of 109,817 PPIs identified in this study, including 1,829 above our CompPASS threshold score.

SUPPLEMENTAL EXPERIMENTAL PROCEDURES

Mitochondrial proteome compilation and disease association

The mouse MitoCarta 2.0 (Mouse.MitoCarta2.0.xls) was downloaded from The Broad Institute (http://www.broadinstitute.org/pubs/MitoCarta/) and Entrez GeneIDs were converted from mouse to human using HomoloGene (build 68) and reciprocal BLASTp searches. Data from the mitochondrial matrix (Rhee et al., 2013) and inter-membrane space (Hung et al., 2014) proteomic studies were then integrated with the human MitoCarta. This list of 1,167 human mitochondrial proteins was used as the basis for subsequent determination of subcellular localization. We further annotated these MXPs by assessing their prior association with disease. Many mitochondrial proteins are associated with disease, often with no clear pathological mechanism (Koopman et al., 2012). To connect mitochondrial proteins with their associated pathology, we mapped large-scale disease lists onto our list of mitochondrial proteins (Amberger et al., 2015; Rath et al., 2012; Safran et al., 2010). We found that 368 mitochondrial proteins are associated with more than 300 human disease states. Notably, despite little to no mechanistic knowledge of their function, we observe that mutations in 29 of our MXPs are causal for disease phenotypes as listed in OMIM and many more have been associated with diseases.

Identification of mitochondrial disease-related genes

Disease-gene annotations were downloaded from OMIM (Amberger et al., 2015), Orphanet (Rath et al., 2012) (downloaded March 2015) and GeneCards/MalaCards (Rappaport et al., 2014; Safran et al., 2010) (Koopman et al., 2012).

Generation of tagged constructs

Mitochondrial open reading frames were obtained from The Broad Institute and DNASU (Seiler et al., 2014). Sequences were compared against UniProt (UniProt, 2015) and dbSNP (Sherry et al., 2001). Any remaining mutations of question were assessed for evolutionary conservation using Homologene. Site-directed mutagenesis was performed by QuikChange PCR and primers were designed by PrimerX software: http://www.bioinformatics.org/primerx/.

Mammalian cell culture

HEK293 and HepG2 cells were grown in Dulbecco's modified Eagle's medium (DMEM, LifeTechnologies) supplemented with 10% fetal bovine serum and 1% penicillin-streptomycin (LifeTechnologies). Cells were subcultured by trypsinization. On day one, 7 million HEK HepG2 cells were plated in a 15 cm dish and allowed to grow overnight. On day two, cells were transiently transfected with a mix of 20 µg pcDNA3.1 gene-FLAG plasmid, 72 µg linear polyethylenimine (PEI, PolySciences), and 900 µL Opti-MEM (LifeTechnologies). On day four, cells were washed with PBS and media was replaced with DMEM containing 10% FBS, 1% penicillin-streptomycin, and either 10 mM glucose or 10 mM galactose. After 24 hours (Day 5), cells were washed with and harvested into phosphate-buffered saline (PBS), were collected at 2,000 rcf, were snap frozen in liquid nitrogen, and were stored at -80°C.

Subcellular localization analysis

On the day prior to transfection, HEK293 cells were plated at a density of 75,000 cells/well onto poly-D-lysine-coated coverslips in 6-well dishes. Cells were transiently transfected with a mix of 1 µg pcDNA3.1 gene-FLAG, 0.5 µg plasmid encoding green fluorescent protein with an N-terminal mitochondrial localization sequence (MLS-GFP) (Hanson et al., 2004), 7.5 µg PEI, or Lipofectamine LTX Plus (Life Technologies), and 200 µL Opti-MEM. After 24 hours, the cells were fixed (4% paraformaldehyde in PBS), permeabilized (0.2% Triton X-100 in PBS), blocked (1% BSA in PBS), and probed with mouse anti-FLAG M2 1° antibody (F1804, Sigma, 1:2000 (v/v) in 1% BSA in PBS) and either Alexa Fluor 594 or 564-conjugated goat anti-mouse 2° antibody (LifeTechnologies, 1:2000 (v/v) in 1% BSA in PBS) in 1% BSA in PBS. Cells were either counter stained with Hoechst dye (1 µg/mL) to label nuclear

DNA and placed in mounting medium (1:1, v/v, glycerol/PBS), or mounted in ProLong Diamond Antifade Mountant with DAPI (Life Technologies). Epifluorescent microscopy was performed on an Olympus IX81 microscope using 100X oil immersion optics.

Affinity enrichment

Digitonin-based approach

Cell pellets were lysed in 200 µL cold lysis buffer (20 mM HEPES, pH 7.40, 100 mM NaCl, 10% glycerol, 3% digitonin (Sigma), 1 mM DTT, protease inhibitors (10 µM benzamide HCl, 1 µg/mL 1,10-phenanthroline and 0.5 µg/mL each of pepstatin A, chymostatin, antipain, leupeptin, aprotinin; Sigma), phosphatase inhibitors (500 µM imidazole, 250 µM NaF, 300 µM sodium molybdate, 250 µM sodium orthovanadate, 1 mM sodium tartrate; Sigma), and deacetylase inhibitors (10 mM each sodium butyrate and nicotinamide; Sigma)), modified from Marbois et al (Marbois et al., 2005). After periodic vortexing on ice, insoluble materials were pelleted (16,000 g, 10 min, 4°C) and the supernatant was retained. The protein concentration was quantified by Bradford assay, and equal masses of cell supernatant were mixed with 30 µL pre-washed anti-FLAG magnetic beads (Sigma M8823) for 2-3 h at 4°C with end-over-end agitation. Following incubation, beads were washed four times in wash buffer (20 mM HEPES, pH 7.40, 100 mM NaCl, 0.05% digitonin, 10% glycerol) and once in final wash buffer (20 mM HEPES, pH 7.40, 100 mM NaCl). Proteins were eluted in 70 µL elution buffer (final wash plus 0.2 mg/mL FLAG-peptide) for 30 min at room temperature with constant agitation.

Quantitative mass spectrometry

Protein digestion

Protein pull-downs (40 µL volume) were transferred from 96-well plates to separate 1.5 mL tubes (Eppendorf) containing 40 µL of 3 M Urea, 100 mM Tris (pH 8). Protein was reduced with 5 mM dithiothreitol (incubation at 37°C for 45 minutes) and alkylated with 15 mM iodoacetamide (incubation in the dark, at ambient temperature, for 45 minutes). Alkylation was quenched by adding an additional 5 mM dithiothreitol (incubation at ambient temperature for 15 minutes). Protein was enzymatically digested with 0.8 µg of sequencing-grade trypsin (Promega, Madison, WI) and incubated at ambient temperature overnight. An additional 0.4 µg of trypsin was added to each sample the next morning and the resulting mixtures were incubated at room temperature for an hour. Digests were quenched by bringing the pH 2 with trifluoroacetic acid and immediately desalted using C18 solid-phase extraction columns (SepPak, Waters, Milford, MA). Prior to washing peptides with 3 mL of 0.1% TFA on the C18 columns, peptides were washed with 1 mL of 5% acetonitrile/0.1% TFA solution to reduce the abundance of the 1x FLAG peptide from each mixture.

CIAF KD cell pellets prepared by pellets were resuspended in 1 mL of 8 M Urea, 40 mM Tris, pH 8, 30 mM sodium chloride, 1 mM calcium chloride and 1 tab protease inhibitor (Roche, Indianapolis, IN). Protein was extracted with 9 mL of methanol (1:10 ratio) and pelleted at 3,500 RPM for 40 minutes. Pellets were air dried for 40 minutes then resuspended in 8 M Urea, 100 mM Tris. Protein concentration was determined through a BCA assay (Thermo Fisher Scientific). Protein (0.5 mg) was reduced with 10 mM TCEP and alkylated with 40 mM chloroacetamide at room temperature. Samples were diluted to 1.5 M Urea with 100 mM Tris, pH 8, and then protein was enzymatically digested with sequencing-grade trypsin in a 1:50 enzyme to protein ratio (Promega, Madison, WI). An additional 1:100 enzyme to protein ratio of sequencing-grade trypsin was added in the morning. Digests were quenched by bringing the pH to 2 with trifluoroacetic acid and immediately desalted using polymeric reversed phase solid-phase extraction columns (StrataX, Phenomenex, Torrance, CA). Peptides were washing with 3 mL of 0.1% TFA and eluted from the columns. Prior to LC-MS/MS analysis, peptide concentration was determined using a quantitative colorimetric peptide assay (Thermo Fisher Scientific).

LC-MS/MS analysis of Protein-Protein interactions

All AE-MS/MS experiments were performed using a NanoAcquity UPLC system (Waters, Milford, MA) coupled to an Orbitrap Elite mass spectrometer (Thermo Fisher Scientific, San Jose, CA). Reverse-phase columns were made

in-house by packing a fused silica capillary (75 µm i.d., 360 µm o.d., with a laser-pulled electrospray tip) with 3.5 µm diameter, 130 Å pore size Bridged Ethylene Hybrid C18 particles (Waters) to a final length of 30 cm. The column was heated to 55°C for all experiments. Samples were loaded onto the column for 12 minutes in 95:5 buffer A [water, 0.2% formic acid, and 5% DMSO] : buffer B [acetonitrile, 0.2% formic acid, and 5% DMSO] at a flow-rate of 0.40 µL/min. Peptides were eluted using the following gradient: an increase to 22% B over 32 min, followed by a 5 min linear gradient from 22% to 28% B, followed by a 3 min linear gradient from 28% to 70% B which was held for 3 minutes. The column was equilibrated with 5% buffer B for an additional 15 min. Precursor peptide cations were generated from the eluent through the utilization of a nanoESI source. Between each set of baits a 3 run wash sequence was utilized to minimize carryover. The first wash consisted of trapping acetonitrile for 25 minutes and the second was cycle consisted of a 45 min high organic gradient. A pre-blank wash was run by injecting 4.5 uL 0.2% formic acid using the 60 minute gradient described above to allow the identification of carryover proteins.

Mass spectrometry instrument methods consisted of MS^1 survey scans ($1x10^6$ target value; 60,000 resolution; 300 Th – 1500 Th) that were used to guide fifteen subsequent data-dependent MS/MS scans (2 Th isolation window, HCD fragmentation, normalized collision energy of 30; $5x10^4$ target value, 15,000 resolution). Dynamic exclusion duration was set to 45 s, with a maximum exclusion list of 500 and an exclusion width of ±10 ppm around the selected average mass. Maximum injection times were set to 50 ms for all MS^1 scans and 200 ms for MS/MS scans.

All CIAF KD experiments were performed on a Dionex UPLC system (Thermo Scientific, San Jose, CA) coupled to an Orbitrap Elite mass spectrometer (Thermo Fisher Scientific, San Jose, CA). Reverse-phase columns were made in-house by packing a fused silica capillary (75 µm i.d., 360 µm o.d., with a laser-pulled electrospray tip) with 1.7 µm diameter, 130 Å pore size Bridged Ethylene Hybrid C18 particles (Waters) to a final length of 30 cm. The column was heated to 65°C.

Samples were loaded onto the column for 6 minutes in 100:0 buffer A [water, 0.2% formic acid, and 5% DMSO] : buffer B [acetonitrile, 0.2% formic acid, and 5% DMSO] at a flow-rate of 0.250 µL/min. Peptides were eluted using the following gradient: an increase to 30% B over 180 min, followed by a 3 min linear gradient from 30% to 100% B, which was held for 3 minutes. The column was equilibrated with 100% buffer A for an additional 20 min. Precursor peptide cations were generated from the eluent through the utilization of a nanoESI source.

Mass spectrometry instrument methods consisted of MS^1 survey scans that were used to guide twenty subsequent data-dependent MS/MS scans. Mass spectrometry instrument methods consisted of MS^1 survey scans ($1x10^6$ target value; 60,000 resolution; 300 Th – 1500 Th) that were used to guide ten subsequent data-dependent MS/MS scans (2 Th isolation window, CAD fragmentation, normalized collision energy of 35; $5x10^4$ target value, ion trap). Dynamic exclusion duration was set to 45 s, with a maximum exclusion list of 500 and an exclusion width of ±10 ppm around the selected average mass. Maximum injection times were set to 50 ms for all MS^1 scans and 75 ms for MS/MS scans.

All raw files have been deposited to the CHORUS Project (http://chorusproject.org), an online repository for sharing, disseminating, and analyzing mass spectrometry data. The AE-MS experiment contains all 936 raw mass spectrometry files from the 6 conditions for each bait, expressed in two different cell lines. The LFQ data contains all 30 files collected from the C17orf89 siRNA lines analyzed. All translated raw files can be viewed directly in a user's web browser or downloaded for sharing, disseminating, and analyzing mass spectrometry data.

Free accounts on Chorus enable users to browse all of the files and download select files or experiments of interest. Otherwise, the full data sets for both categories of experiments can be directly downloaded using the following links:

AE-MS: https://chorusproject.org/anonymous/download/experiment/b4dd75747a60445d9b3d0d16634e794b

LFQ: https://chorusproject.org/anonymous/download/experiment/3dd8e65325734a86b180636f5db2a947

Data Analysis

Data was also processed using the MaxQuant software suite (Cox and Mann, 2008; Cox et al., 2011). Searches were performed against a target-decoy database (UniProt (human) database, www.uniprot.org, April 4th, 2014, containing the 1X FLAG and MFP-GFP-FLAG sequences) using the default settings for high-resolution mass spectra except for FTMS MS/MS tolerance which was set to 0.015 Da. A maximum of 2 missed tryptic cleavages were allowed. The fixed modification specified was carbamidomethylation of cysteine residues. The variable modification specified was oxidation of methionine. Results were filtered to 1% FDR at the unique peptide level and grouped into proteins within MaxQuant. Proteins were quantified across all replicates within each bait set using MaxLFQ (Cox et al., 2014). Matching between runs was allowed with a retention time window of 1 min.

CompPASS (WD) scoring for protein-protein interactions

We scored our interactions based on CompPASS scores as previously described with a two key changes (Behrends et al., 2010): PPI intensity was measured by averaging the Label free quantification (LFQ) intensity and proteins were grouped to avoid punishing interactions observed multiple times within a single complex.

Briefly the LFQ values from the MaxQuant outputs from all groups of 12 replicates (single bait across cell types and carbon source) were combined. WD scores for the interactions were calculated based on the following equations:

$$WD_{i,j} = \sqrt{(\lambda\omega_j)^p (X_{i,j})} \qquad \text{Eq. 1}$$

$$\lambda = \frac{k}{\sum_{i=1}^{i=k} f_{i,j}} \quad , \quad f_{i,j} = \begin{cases} 1; X_{i,j} > 0 \\ X_{i,j} \end{cases} \qquad \text{Eq. 2}$$

$$\omega_j = \frac{\sigma_j}{\overline{X_j}} \quad , \quad \overline{X_j} = \frac{\sum_{i=1}^{i=k} X_{i,j}}{k} \qquad \text{Eq. 3}$$

$$X_{j,i} = \begin{cases} average\ intensity\ of\ the\ interaction \\ between\ bait\ i\ and\ interactor\ j \end{cases}$$

$$p = \begin{cases} number\ of\ replicates \\ runs\ in\ which \\ the\ interactor\ is\ present \end{cases}$$

Where j corresponds to a prey and i corresponds to a bait. k is the total number of different experimental conditions present in the analysis (when all data are used, this number is 78 baits X 2 cell lines X 2 carbon sources = 312.) $f_{i,j}$ indicates if an interaction between a bait (i) and prey (j) is observed ($f_{i,j}$=1) or not ($f_{i,j}$=0.) λ is the reciprocal of frequency indicating the specificity of an interaction, and ω_j is the coefficient of variation for the strength of an interaction (j) across all baits.

A key variation we added to this analysis is the omission of related proteins from the background dataset to improve scoring accuracy for highly interconnected networks. If it is expected that bait A and bait B may be part of the same complex, when λ and ω_j are calculated for bait A, data from bait B is omitted as not to punish interactions from bait A for being present in B (k automatically adjusts for this omission as well). Below is a table of the baits we grouped in this way.

Group Name	Baits
COQ (Lohman et al., 2014)	ADCK3
	ADCK4
	COQ10A
	COQ10B
	COQ2

	COQ3
	COQ4
	COQ5
	COQ6
	COQ7
	COQ9
Complex I (Guarani et al., 2014)	NDUFAF4
	NDUFAF5
	NDUFS3
Ketone Body(Shafqat et al., 2013)	OXCT1
	OXCT2

Protein gels

SDS-PAGE

After competitive elution with FLAG peptide, the eluate was mixed with protein loading buffer and run on a 10% gel for 90 min at 150 V.

Blue Native-PAGE

NativePAGE Novex 4-16% Bis-Tris protein 10-well gels (LifeTechnologies) were prepared with anode buffer (50 mM Bis-Tris) and dark cathode buffer (15 mM Bis-Tris/50 mM tricine, 0.02% Serva Blue G-250). Samples were mixed with NativePAGE sample buffer (LifeTechnologies, BN2003) to 1x final concentration, and were loaded onto gels alongside NativeMARK standard (LifeTechnologies). Gels were run at 150 V for 60 min at 4°C. The dark cathode buffer was then replaced with light cathode buffer (0.002% Serva Blue G-250), and the gels were run for another 60 min at 250 V.

For immunoblotting, gels were transferred to PVDF membranes as below (100 V, 1 h, 4°C). The membrane was then incubated in 8% glacial acetic acid in water for 15 min to fix the proteins. After rinsing with water, methanol was added to the membrane to remove the majority of the Serva Blue stain. Immunoblot proceeded as below.

Immunoblotting

Protein lysates were prepared in RIPA buffer composed of 150 mM NaCl, 1% IGEPAL CA-630, 0.5% sodium deoxycholate, 0.1% SDS, 50 mM Tris (pH 8.0), 0.4 mM EDTA (pH 8.0), 10% glycerol and a protease inhibitor cocktail (Roche Diagnostics), and 20 µg of cleared whole-cell lysate, as determined by BCA assay (Thermo Scientific), was separated on a 4-12% Novex NuPAGE Bis-Tris Mini Gel (Invitrogen), transferred to PVDF and probed with primary antibodies (listed below).

Primary Antibodies used for Immunoblotting

Antigen	Supplier	Product Number
β-actin	Abcam	ab8224
C20orf7 (NDUFAF5)	Abcam	ab192235
CI subunit 8 kDa	Abcam	ab110245
ETFA	Abcam	ab110316
ETFB	Abcam	ab73986
FLAG	Sigma	F1804

NDUFS3	Abcam	ab14711
OxPhos Antibody Cocktail (ATP5A1, SDHA, UQCRC2, NDUFA9, COXIV)	Abcam	ab110412
VDAC	Abcam	ab18988

Generation of C17orf89 knockdown cells (For Figure 3)

Viral particles of the human lentiviral shRNA C17orf89 set (TL320977, OriGene, pGFP-C-shLenti backbone) and nonsilencing lentiviral shRNA controls (TR30021, OriGene; shGFP and empty vector in pLKO.1 backbone, The Broad Institute) were produced in HEK293 cells. The hairpin sequences for the C17orf89 set are: S1: GTTAGAATAAGATGTAACGGAAGCCACGA, S2: GAAGATGTGACATTCCTCGGTGTTAGATC, S3: TCCAGCATTGTGTCCGTAA-ACCTGAGTTA, S4: GGACTCTGAGCTTCACACCTGTCTGCTGC. Cell transfection was carried out with 5.4 µL 10 mM polyethylenimine (PEI) per µg DNA. HEK293T cells (2×10^6) were transfected with 2.5 µg lentiviral shRNA construct together with 1.625 µg psPAX2 and 0.875 µg pMD2.G packaging plasmids. Seventy-two hours after transfection, culture medium containing viral particles was passed through a 0.45 µm filter, was supplemented with 8 µg/ml polybrene, and was added to HEK293 target cells for infection. HEK293 cells were transduced in 6-well plates at 1.5×10^5 cells per well. Transduced HEK293 cells were selected in culture medium supplemented with 2 µg/ml puromycin for at least two weeks. Successful knockdown of C17orf89 was confirmed by quantitative real-time PCR.

Measurement of Complex I activity.

Abcam's Complex I Enzyme Activity Microplate Assay Kit (ab109721) was used to quantify CI activity in C17orf89 knock down cell lines. Normalized protein lysate was loaded onto a CI immunocapture plate and incubated for 3 hours at room temperature. The plate was washed and then NADH + reaction dye mixture was loaded on. The plate was read continuously to measure the rate of NADH turnover.

Measurement of oxygen consumption

All respiration assays were performed using an XF-96 Extracellular Flux Analyzer (Seahorse Bioscience) with a 4-port injection system. For initial comparisons of HEK-293 and HepG2 cells, Seahorse 96-well plates were coated with poly-D-lysine, and trypsinized HEK-293 or HepG2 cells were plated at 25,000 cells per well on the day prior to the assay. After calibration of the assay plate using the XF96 software, cells were loaded into the instrument and subjected to a standard mitochondrial stress test protocol. Data analysis was performed using the XF96 software (version 1.4.1.4), and the "Level(Direct)Akos" algorithm that is a built-in factor of the software package.

To test oxygen consumption rate (OCR) and extracellular acidification rate (ECAR) in C17orf89 knockdown cells, two C17orf89 knockdown cell lines (S2 and S3) and one control line transfected with a non-silencing vector (S5) were plated on poly-D-lysine coated Seahorse 96-well plates at 25,000 cells per well. After a 24-hr incubation, the growth media were exchanged for XF Assay Medium (Seahorse Biosciences) supplemented with 25 mM glucose (Sigma-Aldrich). OCR and ECAR measurements were taken over 3 minute periods following 1 minute mixing periods. After five initial reads, the cells were treated sequentially with FCCP (Cayman Chemical) to a final concentration of 0.1 µM, then rotenone (Sigma-Aldrich) to a final concentration of 1.0 µM, or with equivalent volumes of DMSO as a negative control. Three measurements were taken after each addition.

RNAi For CI assembly factor knock down experiments (for Figure 4)

1 day after seeding, HEK 293 cells were transfected with 10 nM of Dharmacon siGENOME SMARTpools for each target or the Non-Targeting siRNA Control Pool #2 using 0.2% of DharmaFECT 1 Transfection Reagent (GE Healthcare) based on the manufacturer's protocol. After 2 days, the cells were passaged, and the next day were transfected again with 10 nM siRNA, and after another 2 days were collected for real-time qPCR, immunoblot, and mass spectrometry-based proteomic analyses.

Relative Quantification Real Time-qPCR
Total RNA was purified from cultured cells using the RNeasy Mini Kit (QIAGEN). First- strand cDNA was synthesized from purified RNA (500 ng) using the SuperScript III Synthesis System for RT-PCR (Invitrogen). Real time-quantitative PCR was performed using SYBR green-based detection (Applied Biosystems) with *RPLP0* as the endogenous control. Primer sequences are listed below.

Primers used for SYBR Green Real-Time qPCR
SYBR green primers were designed using the Roche Universal ProbeLibrary Assay Design Center tool or the Integrated DNA Technologies (IDT) RealTime qPCR Assay tool.

C17orf89:
Left: ggccgcctgagtaaggac Right: gacaggtgtgaagctcagagtc

NDUFAF4 (C6orf66):
Left: gctcccagacacccctcta Right: cagcttttcatctttacgagca

NDUFAF5 (C20orf7):
Left: gccgaccaaatttgactacc Right: catatacacggtctgcgatcc

NDUFAF6 (C8orf38):
Left: ctggggcactgaccactac Right: agcagcagggagcataaataa

NDUFS3:
Left: tgcacagttcaaatctctggtt Right: cgcagagacaacaggttgtaga

RPLP0:
Left: tctacaaccctgaagtgcttgat Right: caatctgcagacagacactgg

Analysis of CIAF KD lines by mass spectrometry
Protein samples in urea buffer were reduced in 10 mM TCEP and alkylated with 40 mM chloroacetamide at room temperature. Protein was enzymatically digested with sequencing-grade trypsin (Promega, Madison, WI). Digested were quenched by acidification with trifluoroacetic acid and immediately desalted using polymeric reversed phase solid-phase extraction columns (StrataX, Phenomenex, Torrance, CA).
All samples were analyzed using a Dionex UPLC system (Thermo Scientific, San Jose, CA) coupled to an Orbitrap Elite mass spectrometer (Thermo Fisher Scientific, San Jose, CA). Reverse-phase columns were made in-house by packing a fused silica capillary with 1.7 μm diameter, 130 Å pore size Bridged Ethylene Hybrid C18 particles (Waters) to a final length of 30 cm. The column was heated to 65°C. Precursor peptide cations were generated from the eluent through the utilization of a nanoESI source. Mass spectrometry instrument methods consisted of MS1 survey scans that were used to guide 20 subsequent data-dependent MS/MS scans (see Supplemental Experimental Procedures for further details).

Protein digestion
Cell pellets were resuspended in 1 mL of 8 M Urea, 40 mM Tris, pH 8, 30 mM sodium chloride, 1 mM calcium chloride and 1 tab protease inhibitor (Roche, Indianapolis, IN). Protein was extracted with 9 mL of methanol (1:10 ratio) and pelleted at 3,500 RPM for 40 minutes. Pellets were air dried for 40 minutes then resuspended in 8 M Urea, 100 mM Tris. Protein concentration was determined through a BCA assay (Thermo Fisher Scientific). Protein (0.5 mg) was reduced with 10 mM TCEP and alkylated with 40 mM chloroacetamide at room temperature. Samples

were diluted to 1.5 M Urea with 100 mM Tris, pH 8, and then protein was enzymatically digested with sequencing-grade trypsin in a 1:50 enzyme to protein ratio (Promega, Madison, WI). An additional 1:100 enzyme to protein ratio of sequencing-grade trypsin was added in the morning. Digests were quenched by bringing the pH to 2 with trifluoroacetic acid and immediately desalted using polymeric reversed phase solid-phase extraction columns (StrataX, Phenomenex, Torrance, CA). Peptides were washing with 3 mL of 0.1% TFA and eluted from the columns. Prior to LC-MS/MS analysis, peptide concentration was determined using a quantitative colorimetric peptide assay (Thermo Fisher Scientific).

LC-MS/MS Analysis

All samples were analyzed using a Dionex UPLC system (Thermo Scientific, San Jose, CA) coupled to an Orbitrap Elite mass spectrometer (Thermo Fisher Scientific, San Jose, CA). Reverse-phase columns were made in-house by packing a fused silica capillary (75 µm i.d., 360 µm o.d., with a laser-pulled electrospray tip) with 1.7 µm diameter, 130 Å pore size Bridged Ethylene Hybrid C18 particles (Waters) to a final length of 30 cm. The column was heated to 65°C.

Samples were loaded onto the column for 6 minutes in 100:0 buffer A [water, 0.2% formic acid, and 5% DMSO] : buffer B [acetonitrile, 0.2% formic acid, and 5% DMSO] at a flow-rate of 0.250 µL/min. Peptides were eluted using the following gradient: an increase to 30% B over 180 min, followed by a 3 min linear gradient from 30% to 100% B, which was held for 3 minutes. The column was equilibrated with 100% buffer A for an additional 20 min. Precursor peptide cations were generated from the eluent through the utilization of a nanoESI source.

Mass spectrometry instrument methods consisted of MS1 survey scans that were used to guide twenty subsequent data-dependent MS/MS scans. Mass spectrometry instrument methods consisted of MS1 survey scans (1x106 target value; 60,000 resolution; 300 Th – 1500 Th) that were used to guide ten subsequent data-dependent MS/MS scans (2 Th isolation window, CAD fragmentation, normalized collision energy of 35; 5x104 target value, ion trap). Dynamic exclusion duration was set to 45 s, with a maximum exclusion list of 500 and an exclusion width of ±10 ppm around the selected average mass. Maximum injection times were set to 50 ms for all MS1 scans and 75 ms for MS/MS scans.

Coenzyme Q Quantification

A frozen pellet of tissue culture cells (isolated from a 20 cm plate at 80–100% confluence) was thawed on ice, mixed with glass beads (0.5 mm diameter, 100 µL) and H_2O (400 µL, 4 °C), and vortexed (4 x 30 s) to lyse the cells. A portion of the cell lysate (containing 2 mg of protein, as determined by a BCA assay) was mixed with H_2O (to a total volume of 200 µL), spiked with an internal standard (CoQ_6, 20 µL, 10 µM), and mixed by vortexing (30 s). $CHCl_3$/MeOH (1:1, v/v) (900 µL) was added and vortexed (3 x 30 s). The samples were centrifuged (3,000 *g*, 3 min, 4 °C) to complete phase separation. 600 µL of the lower organic phase was transferred to a clean tube. HCl (1 M, 100 µL, 4 °C) was added to the remaining aqueous phase and vortexed (30 s). $CHCl_3$/MeOH (1:1, v/v) (400 µL) was added and vortexed (2 x 30 s). The samples were centrifuged (3,000 *g*, 3 min, 4 °C) to complete phase separation. The lower organic phase was combined with the previous organic extract, and the combined extracts were dried under $Ar_{(g)}$. The organic residue was reconstituted in ACN/IPA/H_2O (65:30:5, v/v/v) (100 µL) by vortexing (2 x 30 s).

Lipids from 10 µL of extract were separated by LC on an Ascentis Express C18 column (150 mm x 2.1 mm x 2.7 µm particle size, Supelco, Bellefonte, PA) using an Accela HPLC pump (Thermo Scientific, San Jose, CA) at a flow-rate of 0.5 mL/min using a linear gradient. Mobile phase A was 70/30 acetonitrile/water containing 10 mM ammonium acetate and 0.025% acetic acid and B was 90/10 isopropanol/acetonitrile containing the same additives. Initially the flow was maintained at 50% B for 1.5 min, then ramped to 95% B over 6.5 min, held there for 2 min before returning to starting conditions over 1 min, and finally re-equilibrating the column for 2.5 min. The auto sampler (HTC PAL, Thermo Scientific) vigorously mixed each sample before injection to ensure homogeneity. Quantitation was performed using tandem mass spectrometry by integrating the peak area of the 880.72 à 197.08 Th transition of coenzyme Q_{10} and normalizing to the 591.44 à 197.08 Th transition of the internal standard coenzyme Q_6 using the Xcalibur software suite (Thermo Scientific, Version 3.0). MS conditions were as follows: a Q Exactive

mass spectrometer (Thermo Scientific, Build 2.5) equipped with a HESI II spray source kept at 350 °C and +4 kV was used for detection. The inlet capillary was kept at 350 °C, sheath gas was set to 60 units, and auxiliary gas to 20 units. Resolving power was set to 35,000, AGC target to $2x10^5$, isolation width to 1 Th, normalized collision energy to 27 units, and maximum injection time to 375 ms.

Cell-free expression and purification.

Plasmid DNA was treated with 0.05 μg/μL proteinase K to remove trace amounts of RNase, purified, and used as individual transcription templates with SP6 RNase polymerase. Transcription and translation methods are as previously described (Makino et al., 2013). Briefly, transcription reactions included 0.2 mg/mL DNA, 20 mM magnesium acetate, 2 mM spermidine trihydrochloride, 10 mM DTT, 80 mM Hepes-KOH, pH 7.8, 4 mM each NTP, pH 7.0, 1.6 U/μL SP6 RNA polymerase (Promega), and 1 U/μL RNasin (Promega). Transcriptions were incubated for 4 h at 37°C. Nonpurified transcription reactions were used singly or mixed in equal volumes for multi-protein interaction monitoring and added to wheat germ extract (WEPRO2240; CellFree Sciences) in a standard dialysis cup reaction (Makino et al., 2013). Each 25 μL reaction contained 60 O.D. wheat germ extract (6.25 μL), 24 mM Hepes-KOH, pH 7.8, 100 mM potassium acetate, 6.25 mM magnesium acetate, 0.4 mM spermidine trihydrochloride, 4 mM DTT, 1.2 mM ATP, 0.25 mM GTP, 16 mM creatine phosphate, 0.0005% sodium azide, 0.04 mg/mL creatine kinase, 0.3 mM each amino acid (amino acid mixture adjusted to pH 7.0 with KOH), and 5 μL of RNA. Thirty-two-fold excess dialysis buffer was used for each reaction, containing all reaction components except for wheat germ extract, creatine kinase, and RNA. After assembly of dialysis cups and 18-h incubation at 23°C , duplicate translations (50 μL) of each combination were pooled and centrifuged for 5 min at 20,000 rcf at 10°C. The soluble fraction was removed and added to 20 μL of StrepTactin resin equilibrated in 25 mM Hepes, pH 7.8, 150 mM NaCl, and 1 mM DTT in a 96-well filter plate (HTS Multiscreen; Millipore). The resin with bound protein was washed three times with 150 μL of binding buffer. Strep(II)-tagged protein and protein complexes were eluted in 2.5 mM desthiobiotin in the same buffer. Bound samples were eluted by increasing the imidazole concentration to 500 mM. Samples of the StrepTactin elutions were loaded without heating on 4-20% Stain-free TGX Criterion gels (Bio-Rad) and subjected to denaturing SDS/PAGE. Gels were imaged by tryptophan fluorescence using a Bio-Rad Stain Free Imager, followed by staining in Coomassie Brilliant Blue R-250.

SUPPLMENTAL REFERENCES

Amberger, J.S., Bocchini, C.A., Schiettecatte, F., Scott, A.F., and Hamosh, A. (2015). OMIM.org: Online Mendelian Inheritance in Man (OMIM®), an online catalog of human genes and genetic disorders. Nucleic Acids Research *43*, D789-798.

Behrends, C., Sowa, M., Gygi, S., and Harper, J. (2010). Network organization of the human autophagy system. Nature *466*, 68-U84.

Cox, J., Hein, M.Y., Luber, C.A., Paron, I., Nagaraj, N., and Mann, M. (2014). Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ. Mol Cell Proteomics *13*, 2513-2526.

Cox, J., and Mann, M. (2008). MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. Nature biotechnology *26*, 1367-1372.

Cox, J., Neuhauser, N., Michalski, A., Scheltema, R.A., Olsen, J.V., and Mann, M. (2011). Andromeda: a peptide search engine integrated into the MaxQuant environment. Journal of proteome research *10*, 1794-1805.

Guarani, V., Paulo, J., Zhai, B., Huttlin, E.L., Gygi, S.P., and Harper, J.W. (2014). TIMMDC1/C3orf1 functions as a membrane-embedded mitochondrial complex I assembly factor through association with the MCIA complex. Mol Cell Biol *34*, 847-861.

Hanson, G.T., Aggeler, R., Oglesbee, D., Cannon, M., Capaldi, R.A., Tsien, R.Y., and Remington, S.J. (2004). Investigating mitochondrial redox potential with redox-sensitive green fluorescent protein indicators. The Journal of biological chemistry *279*, 13044-13053.

Hung, V., Zou, P., Rhee, H.-W., Udeshi, N.D., Cracan, V., Svinkina, T., Carr, S.A., Mootha, V.K., and Ting, A.Y. (2014). Proteomic Mapping of the Human Mitochondrial Intermembrane Space in Live Cells via Ratiometric APEX Tagging. Molecular cell.

Koopman, W.J.H., Willems, P.H.G.M., and Smeitink, J.A.M. (2012). Monogenic Mitochondrial Disorders. New England Journal of Medicine *366*, 1132-1141.

Lohman, D.C., Forouhar, F., Beebe, E.T., Stefely, M.S., Minogue, C.E., Ulbrich, A., Stefely, J.A., Sukumar, S., Luna-Sanchez, M., Jochem, A.*, et al.* (2014). Mitochondrial COQ9 is a lipid-binding protein that associates with COQ7 to enable coenzyme Q biosynthesis. Proc Natl Acad Sci U S A *111*, E4697-4705.

Makino, S.-i., Beebe, E.T., Markley, J.L., and Fox, B.G. (2013). Cell-Free Protein Synthesis for Functional and Structural Studies. In (Totowa, NJ: Humana Press), pp. 161-178.

Marbois, B., Gin, P., Faull, K.F., Poon, W.W., Lee, P.T., Strahan, J., Shepherd, J.N., and Clarke, C.F. (2005). Coq3 and Coq4 define a polypeptide complex in yeast mitochondria for the biosynthesis of coenzyme Q. J Biol Chem *280*, 20231-20238.

Rappaport, N., Twik, M., Nativ, N., Stelzer, G., Bahir, I., Stein, T.I., Safran, M., and Lancet, D. (2014). MalaCards: A Comprehensive Automatically-Mined Database of Human Diseases. Current protocols in bioinformatics *47*, 1.24.21-21.24.19.

Rath, A., Olry, A., Dhombres, F., Brandt, M.M., Urbero, B., and Ayme, S. (2012). Representation of rare diseases in health information systems: the Orphanet approach to serve a wide range of end users. Human mutation *33*, 803-808.

Rhee, H.-W., Zou, P., Udeshi, N.D., Martell, J.D., Mootha, V.K., Carr, S.A., and Ting, A.Y. (2013). Proteomic mapping of mitochondria in living cells via spatially restricted enzymatic tagging. Science (New York, NY) *339*, 1328-1331.

Safran, M., Dalah, I., Alexander, J., Rosen, N., Iny Stein, T., Shmoish, M., Nativ, N., Bahir, I., Doniger, T., Krug, H.*, et al.* (2010). GeneCards Version 3: the human gene integrator. Database *2010*, baq020-baq020.

Seiler, C.Y., Park, J.G., Sharma, A., Hunter, P., Surapaneni, P., Sedillo, C., Field, J., Algar, R., Price, A., Steel, J.*, et al.* (2014). DNASU plasmid and PSI:Biology-Materials repositories: resources to accelerate biological research. Nucleic Acids Research *42*, D1253-1260.

Shafqat, N., Kavanagh, K.L., Sass, J.O., Christensen, E., Fukao, T., Lee, W.H., Oppermann, U., and Yue, W.W. (2013). A structural mapping of mutations causing succinyl-CoA:3-ketoacid CoA transferase (SCOT) deficiency. J Inherit Metab Dis *36*, 983-987.

Sherry, S.T., Ward, M.H., Kholodov, M., Baker, J., Phan, L., Smigielski, E.M., and Sirotkin, K. (2001). dbSNP: the NCBI database of genetic variation. Nucleic Acids Research *29*, 308-311.

UniProt, C. (2015). UniProt: a hub for protein information. Nucleic Acids Res *43*, D204-212.