Mitochondrial peptide BRAWNIN is essential for vertebrate respiratory complex III assembly

Shan Zhang, Boris Reljić, Chao Liang, Baptiste Kerouanton, Joel Celio Francisco, Jih Hou Peh, Camille Mary, Narendra Suhas Jagannathan, Volodimir Olexiouk, Claire Tang, Gio Fidelito, Srikanth Nama, Ruey-Kuang Cheng, Caroline Lei Wee, Loo Chien Wang, Paula Duek Roggli, Prabha Sampath, Lydie Lane, Enrico Petretto, Radoslaw M. Sobotka, Suresh Jesuthasan, Lisa Tucker-Kellogg, Bruno Reversade, Gerben Menschaert, Lei Sun, David A. Stroud & Lena Ho

The emergence of small open reading frame (sORF)-encoded peptides (SEPs) is rapidly expanding the known proteome at the lower end of the size distribution. Here, we show that the mitochondrial proteome, particularly the respiratory chain, is enriched for small proteins. Using a prediction and validation pipeline for SEPs, we report the discovery of 16 endogenous nuclear encoded, mitochondrial-localized SEPs (mito-SEPs). Through functional prediction, proteomics, metabolomics and metabolic flux modeling, we demonstrate that BRAWNIN, a 71 a.a. peptide encoded by C12orf73, is essential for respiratory chain complex III (CIII) assembly. In human cells, BRAWNIN is induced by the energy-sensing AMPK pathway, and its depletion impairs mitochondrial ATP production. In zebrafish, Brawnin deletion causes complete CIII loss, resulting in severe growth retardation, lactic acidosis and early death. Our findings demonstrate that BRAWNIN is essential for vertebrate oxidative phosphorylation. We propose that mito-SEPs are an untapped resource for essential regulators of oxidative metabolism.
The size and complexity of the human proteome is constantly evolving, ranging from genome-based estimations of 20,000 canonical proteins to over 70,000 if one considers protein variants (Uniprot, 2020). With the advent of ribosome profiling, the question of the size of the human proteome must once again be revisited. By detecting RNA translation at sub-codon resolution1, ribosome profiling has unexpectedly uncovered thousands of coding regions hitherto thought to be translationally silent2. In particular, small open reading frames (sORFs) nestled within long non-coding RNAs (lncRNAs), pseudogenes, intergenic transcripts or upstream of known ORFs (uORFs) that were previously not considered to be functional have now been shown to encode bona fide peptides3–5. Evidence accumulating within the last decade point to the widespread prevalence of sORF-encoded peptides, or SEPs, and have spurred multiple independent efforts to discover and catalog them6–8. From a biological point of view, the SEP mini-proteome is a rich repository of unexplored gene functions. However, unlike the high throughput techniques that enabled their discovery, functional de-orphanization of SEPs remains a resource-intensive endeavor that relies on fundamental approaches in experimental biology. To date, SEPs have been implicated in diverse functions ranging from cardiovascular and placental development, mTORC1 activation, calcium signaling, mRNA de-capping to membrane fusion, and oocyte fertilization9–15. However, whether SEPs have programmatic functions enabled by their small size remains an open question. To this end, we asked if SEPs were particularly enriched in subcellular locations that would provide instructive cues for their functional characterization.

Results

Mitochondria proteome is enriched in SEPs. To this point, we observed that the mitochondrial proteome, particularly at the inner mitochondrial membrane (IMM) (cataloged in Mitocarta 2.016) is enriched for proteins smaller than 100 a.a. i.e. SEPs (p-value = 2.2×10−16) (Fig. 1a–c), and speculated that mitochondria are a hotspot for SEP function. To uncover novel mitochondria-targeted SEPs (mito-SEPs) from the human SEP peptidome9, 2343 high-quality human SEPs encoded by annotated short-coding sequence (CDS) genes, lncRNAs, uORFs, and intergenic ORFs were selected for test of mitochondrial localization (Fig. 1d, details in Supplementary Note 1). We first employed a three-prong strategy that identified a candidate mito-SEP on the basis of a mitochondrial gene expression signature (Supplementary Fig. 1a); a mitochondrial targeting protein domain (Supplementary Fig. 1b) or empirical detection of the SEP in mass spectrometric data of purified mitochondria (details in Supplementary Note 1). Altogether, our mito-SEP discovery pipeline identified 173 candidates (Fig. 1d and Supplementary Data 1), including 2 positive controls MKKS uORF1 and uORF217. HA-tagged mito-SEP candidates were experimentally validated by immunofluorescence of transfected HeLa cells (Fig. 1e). In total, we successfully validated protein expression for 88/173 sORFs (Supplementary Data 1), defined as the ability of the ORF to generate a peptide with sufficient stability to enable detection (Supplementary Fig. 1c–e). Of these, 23% (20/88) of successfully expressed peptides displayed unambiguous mitochondrial localization including the MKKS uORFs (Fig. 1e). Among the positive candidates were MTLN/MOXI and MIEF1 uORF which were subsequently independently discovered to have critical roles in mitochondria metabolism and mtDNA translation18–21, providing proof for the robustness of our pipeline. All in all, our screen identified 16 bona fide mito-SEPs of unknown function (MSUP) encoded by uORFs, lncRNAs and short CDSs (Fig. 1e and Table 1).

BRAWNIN resides in the IMM and responds to AMPK. We selected a mito-SEP encoded by the C12orf73 gene, which we renamed BRAWNIN (BR) (Fig. 2a and Supplementary Fig. 2a), for further characterization because its stable overexpression in U87MG significantly enhanced OXPHOS (Supplementary Fig. 2b). The BR peptide is conserved in vertebrates (Fig. 2b) but not in yeast and nematodes. Although ribosome profiling detected 2 plausible BR peptides arising from alternative splicing (Fig. 2a and Supplementary Fig. 2c), we confirmed that only the cDNA of the long isoform (P1) produced a stable 8-kDa peptide, while the short isoform (P2) was remarkably less stable even when over-expressed (Supplementary Fig. 2d). Endogenous BR was detected by western blotting with a custom α-BR antibody in HEK293T and HeLa cell lysates (Fig. 2c) and can be depleted by siRNAs targeting the BR transcript (Fig. 2d, e). Endogenous BR is enriched in mitochondria-rich fractions (Fig. 2f), and co-localized with Mitotracker in both HeLa and HEK293T (Fig. 2g, h). Altogether, these data confirm that BR is a nuclear-encoded peptide that is imported into the mitochondria. In vivo, BR can be detected in mouse skeletal muscle where it co-localizes with matrix marker Citrate Synthase (CS) and with a mitochondria-targeted GFP transgene mito-Dendra22 (Fig. 2i, j). BR is also detectable in human cardiac and skeletal muscle (Supplementary Fig. 2e) where it displays a staining pattern characteristic of the mitochondria network. Moreover, BR protein abundance in mouse tissues correlated well with that of mitochondrial respiratory chain proteins, being high in brown adipose, cardiac and skeletal muscle but virtually undetectable in white adipose tissue (Supplementary Fig. 2f). Together, these results establish that BR is a bona fide endogenous mito-SEP both in vitro and in vivo.

Because mitochondria are highly compartmentalized, the molecular function of a protein determines its sub-mitochondrial location and vice versa24. We first established that BR is membrane-bound by sodium carbonate extraction (Fig. 2k). Next, using differential detergent extraction, we determined that BR resides in the IMM and not in the outer mitochondrial membrane (OMM) (Fig. 2l). Proteinase K protection assays further suggest that the C-terminus of BR faces the intermembrane space (IMS) (Fig. 2m). In contrast to its clear mitochondrial localization, human BR lacks a classical mitochondrial targeting sequence (MTS); its N-terminus is instead predicted to function as a secretory signal peptide by SignalP25. Indeed, when overexpressed, BR can be detected in the supernatant (Supplementary Fig. 2g). However, this may not represent bona fide secretion as it cannot be inhibited by brefeldin A. Hence, the N-terminal “signal peptide”, which is not cleaved (Supplementary Fig. 2h), functions as a mitochondrial targeting single-pass transmembrane domain (TMD) to anchor endogenous BR into respiratory complexes. Mitochondria are the principal sites of energy conversion where ATP is generated through the process of oxidative phosphorylation (OXPHOS). Defects in OXPHOS are responsible for many inherited mitochondrial diseases and manifest in aging and metabolic disorders such as diabetes, heart disease, and cancer22. Notably, 63.5% (40/63) of known mitochondrial low molecular weight (MW) proteins function as assembly factors or core sub-units of respiratory chain (RC) complexes that carry out OXPHOS (Supplementary Fig. 1f). RC complexes are vice versa enriched for low MW proteins (Supplementary Fig. 1g). Of the 16 mito-SEPs, the expression of 8 correlated with respiratory chain and electron transport function above the 90th percentile (Fig. 1f), suggesting that half of them participate directly or indirectly in oxidative phosphorylation.
the IMM. Consistent with this, BR-P2, which contains the full TMD of BR-P1, is similarly targeted to the mitochondria despite its relative instability (Supplementary Fig. 2i). Removal of the predicted TMD from BR (BRΔN25), while causing its near complete destabilization (Supplementary Fig. 2j), targets BRΔN25 to the cytosol instead of the mitochondria (Supplementary Fig. 2k).

Proteins in the IMM participate in a wide variety of processes, most commonly OXPHOS, calcium import, and the transport of proteins and metabolites. Co-expression analysis by weighted gene co-expression network analysis (WGCNA)26 supports the notion that BR participates in electron transport and OXPHOS in multiple human organs and species (Fig. 3a, Supplementary Fig. 3a, Supplementary Data 2). Components of the respiratory
chain are dynamically regulated by energy status. In response to a reduction of cellular ATP:ADP ratio, such as during nutrient depletion or exercise, the energy-sensing AMP-activated protein kinase (AMPK) pathway restores energetic balance by stimulating ATP-consuming anabolic processes. As such, we hypothesized that BR levels would be regulated by cellular energy status. Indeed, activation of AMPK using 5-aminoimidizole-4-carboxamide-1-beta-d-ribose (AICAR) robustly increased BR production. To investigate this, we silenced BR in glioblastoma cells (a highly oxidative cell type) using both shRNA and siRNA approaches (Supplementary Fig. 3e). These data confirm that BR is essential for cellular bioenergetics, potentially through regulation of the electron transport chain.

Deletion of Br in zebrafish causes mitochondrial disease. BR is highly conserved in vertebrates, suggesting that its role in bioenergetics is fundamental and required throughout evolution. To understand the in vivo requirement of BR in animal physiology, we turned to the zebrafish where br is most highly expressed in skeletal and cardiac muscle (Fig. 4a). Using Crispr/Cas9, we generated protein null br knockout animals (KO) by removing the entire protein-coding ORF contained in exons 2 and 3 (Supplementary Fig. 4a and Fig. 4b). br zygotic KO (zKO) larvae were produced at Mendelian ratios (Fig. 4c) with no gross body size by 65 dpf (Fig. 4e). Juveniles survived into adulthood, but displayed growth retardation (Fig. 4d). However, if left unsegregated, br zKOs could not be recovered by 65 days post fertilization (dpf) (Fig. 4c). We therefore segregated larvae from intercrosses at 3 dpf and allowed them to grow at low density. Under this regime, zKOs juveniles survived into adulthood, but displayed growth retardation from 28 dpf (Fig. 4d) that resulted in large differences in body size by 65 dpf (Fig. 4e). zKOs are sub-fertile, and can generate maternal zygotic knockout embryos (mzKOs) albeit at a very low frequency (~1/20 mating pairs). Because Br is a mitochondrial peptide, it is maternally deposited in oocytes. mzKOs therefore lack any residual Br protein that is available to support mitochondrial function.

Table 1 List of verified mito-SEPs.

| Label   | Name    | Gene ID                  | Locus | Motif | Length (a.a.) |
|---------|---------|--------------------------|-------|-------|---------------|
| BRAWNN1 | BRAWNN1 | ENSG00000204954          | sORF  | SP    | 71            |
| MSUFL1  | PIGBOS1 | ENSG000000225973         | sORF  | TMD   | 54            |
| MSUFL2  | LINCO1960 | ENSG000000196741       | sORF  | TMD   | 54            |
| MSUFL3  | LINCO1973 | ENSG000000223414       | Incrna| TMD   | 94            |
| MSUFL4  | SIMM192 | ENSG000000163866         | sORF  | TMD   | 92            |
| MSUFL5  | SIMM292 | ENSG000000232388         | sORF  | TMD   | 95            |
| MSUFL6  | C16ORF91 | ENSG00000171409         | sORF  | TMD   | 65            |
| MSUFL7  | PRPF91 uORF | ENSG00000110107    | SUTR  | MTS   | 66            |
| MSUFL8  | UBQ1N1 uORF | ENSG00000135018   | SUTR  | MTS   | 53            |
| MSUFL9  | RNASEH2C uORF | ENSG00000172922 | SUTR  | MTS   | 79            |
| MSUFL10 | CENPW uORF | ENSG00000203760       | SUTR  | MTS   | 68            |
| MSUFL11 | MRPL24 uORF | ENSG00000143314      | SUTR  | Not detected | 77          |
| MSUFL12 | TMEM222 uORF | ENSG00000186501   | SUTR  | TMD   | 90            |
| MSUFL13 | EIF3K uORF | ENSG00000178982       | SUTR  | Not detected | 59          |
| MSUFL14 | KIAA0100 uORF | ENSG0000007202   | SUTR  | TMD   | 73            |
| MSUFL15 | SLC35A4 uORF | ENSG00000176087  | SUTR  | TMD   | 96            |
| MTLN    | LINCO116 | ENSG00000175701        | sORF  | TMD   | 56            |
| MIEF1 uORF | MIEF1 uORF | ENSG00000285025 | SUTR  | MTS   | 70            |

SP signal peptide, TMD transmembrane domain, MTS mitochondrial targeting sequence.
Proteinase K protection assay of HEK293T mitochondria. The digitonin. TOMM20, TOMM70, and VDAC1 are outer mitochondrial membrane (OMM) proteins, are MTCO1, ATP5A, and TIM23 are IMM associated. Both CS and ATP5A are in the matrix. TIM23 has IMS domains. Both CS and ATP5A are in the matrix.

Western blot of HEK293T lysates transfected with control or BR-targeting siRNAs. Scale = 5 μM. j α-BR and α-citrate synthase (CS) immunofluorescence in mouse skeletal muscle, longitudinal section. Scale = 20 μM. k Extraction of endogenous BR from HEK293T mitochondria with buffer alone, Na2CO3 solution or buffer containing 1% Triton X. S = supernatant; P = pellet. l Extraction of endogenous BR from HEK293T mitochondria using increasing concentrations of digitonin. TOMM20, TOMM70, and VDAC1 are outer mitochondrial membrane (OMM) proteins, are MTCO1, ATP5A, and TIM23 are IMM associated. m Proteinase K protection assay of HEK293T mitochondria. The α-BR epitope is C-terminal to the predicted transmembrane domain. TOMM20 and TOMM70 have cytosolic domains. TIM23 has IMS domains. Both CS and ATP5A are in the matrix.

Fig. 2 BRAWNIN (BR) is a conserved SEP at the inner mitochondrial membrane. a Ribosome protected fragment (RPF) reads from the BRAWNIN gene (CT2or73) in human cell lines. P1 and P2 refer to the 2 potential ORF isoforms generated by alternative splicing. b BR (P1) peptide is conserved throughout vertebrate evolution and possesses an N-terminal hydrophobic region that is predicted to be a transmembrane domain or signal peptide. Underlined regions are peptides detected by mass spec. Regions used as immunogens for raising the polyclonal antibodies used in this study are indicated. c Endogenous BR is detected in HeLa and HEK293T with an α-BR antibody. OE = overexpression. d Western blot of HEK293T lysates transfected with control or BR-targeting siRNAs. e Immunofluorescence of endogenous BR detected by α-BR in HeLa transfected with control or BR-targeting siRNAs. Scale = 5 μM. f Western blot of indicated proteins from subcellular fractions of HEK293T. W = whole cell, M = mitochondria enriched fraction, C = cytosol, P = nuclear fraction and debris, including un-extracted mitochondria; BRΔN25 = synthetic BR peptide lacking the N-terminal 25 a.a. g Endogenous BR detected by α-BR in HeLa costained with mitotracker. Scale = 5 μM. h Endogenous BR detected by α-BR in HEK293T costained with mitotracker. Scale = 5 μM. i α-BR and α-citrate synthase (CS) immunofluorescence in mouse skeletal muscle, longitudinal section. Scale = 20 μM. j α-BR immunofluorescence and GFP epitope in mouse skeletal muscle, longitudinal section. Scale = 20 μM. k Extraction of endogenous BR from HEK293T mitochondria with buffer alone, Na2CO3 solution or buffer containing 1% Triton X. S = supernatant; P = pellet. l Extraction of endogenous BR from HEK293T mitochondria using increasing concentrations of digitonin. TOMM20, TOMM70, and VDAC1 are outer mitochondrial membrane (OMM) proteins, are MTCO1, ATP5A, and TIM23 are IMM associated.

m Proteinase K protection assay of HEK293T mitochondria. The α-BR epitope is C-terminal to the predicted transmembrane domain. TOMM20 and TOMM70 have cytosolic domains. TIM23 has IMS domains. Both CS and ATP5A are in the matrix.
had a more severe phenotype compared to zKOs (Fig. 4d). By 6 dpf, mzKO larvae were visibly less active despite fully inflated swim bladders (Supplementary Fig. 4e). By 7 dpf, at the onset of feeding, their food intake was significantly less than WT larvae (Supplementary Fig. 4f). By 11 dpf, they displayed a dramatic reduction of directional swimming and distance swam compared to WTs (Fig. 4f and Supplementary Fig. 4g). Consequently, mzKO larvae did not increase in body length even under low density segregated growth conditions, and did not survive past 42 dpf under the conditions of our aquatics facility (Fig. 4d). In line with our observations in human U87MG cells, basal, ADP stimulated, maximal respiration, and ATP production in br KO skeletal muscle mitochondria were reduced to half that of WT (Fig. 4g). This in turn significantly reduced whole animal respiration br mz KO compared to WT larvae even as early as 1 dpf (Fig. 4h), leading to a failure to thrive and early death (Fig. 4d). These data provide evidence that BR is mito-SEP with conserved and critical functions in mediating OXPHOS, and in vivo deletion of br causes severe mitochondrial deficiency leading to growth retardation.

**BRAWNIN is required for respiratory chain complex III.** To understand the molecular basis of the mitochondrial defect in br KO fish, we profiled the metabolomic perturbations in whole mzKO larvae at 5 dpf. Targeted metabolomics analysis revealed a 2-fold increase in the levels of lactate in mzKO larvae (Fig. 5a) and a 3-fold increase in the citric acid cycle intermediate succinate as compared to WT larvae (Fig. 5b). Lactate accumulation is caused by decreased pyruvate oxidation secondary to mitochondrial deficiency and is frequently used as a diagnostic symptom for patients with mitochondrial diseases. To understand what molecular defects in the mitochondria could result in these perturbations, we turned to MitoCore, a metabolic flux model that simulates over 300 mitochondrial reactions in a human cardiomyocyte using flux balance analysis. By blocking each reaction in turn, and simulating 5000 feasible flux states for each possible blockage, we found that inhibiting ETC Complex III (CIII, Fig. 5c and Supplementary Fig. 5a) and Complex IV (CIV, Supplementary Fig. 5b) flux caused the largest, most biologically plausible increase in succinate and lactate (as inferred from the export rate of the two metabolites). Similarly, when we deleted each of the 371 genes in the MitoCore model, the deletion of genes encoding CIII and CIV components caused the largest increase in succinate and lactate (Supplementary Table 1). In contrast, inhibiting Complex II was not sufficient to induce the observed elevations in lactate (Supplementary Fig. 5c). These data suggest that the underlying molecular defect in br KO mitochondria is related to CIII and/or CIV dysfunction.

To investigate this possibility, we constructed the BR interactor in HEK293T mitochondria by co-immunoprecipitating...
Among all genes encoding RC proteins, BR transcript expression correlates most strongly with those encoding CIII assembly factors (Supplementary Data 4). On this basis, we hypothesized that the observed respiratory defects in si/shBR U87MG and br KO fish are caused by CIII deficiency. To address this possibility, we performed quantitative proteomics on isolated skeletal muscle mitochondria from br WT and KO animals (Fig. 6a). Strikingly, we found that 5 subunits of CIII were significantly downregulated, while subunits of CI, II and IV were not significantly affected (Fig. 6b, c, Supplementary Data 5), suggesting that CIII stability or assembly is specifically impaired in the absence of br. Indeed, SDS-PAGE analysis of br KO mitochondria confirmed that total Uqcr1c and Uqrc2 levels were reduced by >80% (Fig. 6d, e) resulting in a >90 % reduction of mature CIII dimers (CIII2) as measured by Uqcr1c, Uqrc2, Uqcrh and Uqcrb densitometry in blue-native (BN)-PAGE (Fig. 6f). The specific reduction in CIII2 was also seen in shBR U87MG cells both at steady state (Supplementary Fig. 6a–c) and during active assembly of respiratory chain complexes released from doxycycline-mediated inhibition of mitochondrial translation (Supplementary Fig. 6d). Furthermore, deletion of murine Br in mouse embryonic fibroblasts (MEF) (Supplementary Fig. 6e, f) likewise resulted in reduction of CIII subunit proteins levels and assembled CIIIls (Supplementary Fig. 6g, h), demonstrating that the requirement for BR in CIII assembly or stability extends to either endogenous or overexpressed BR followed by mass spectrometry (IP/MS). At the intersection of 4 independent IP/MS experiments were three proteins: BR, GHITM, and UQCRC1 (Fig. 5d, Supplementary Data 3). UQCRC1 is a core component of CIII33 while GHITM has been proposed to regulate cytochrome c release during apoptosis34. The interaction between BR and UQCRC1 detected in human cells was validated by co-immunoprecipitation (co-IP)/western blot in digitonin-solubilized mouse heart mitochondria expressing a Br-FLAG construct (Fig. 5e). Likewise, BR was co-immunoprecipitated by CIII subunit UQCRQ using a surface exposed C-terminal FLAG tag in mouse embryonic fibroblasts (MEF) (Supplementary Data 4). On this basis, we hypothesized that the observed respiratory defects in si/shBR U87MG and br KO fish are caused by CIII deficiency. To address
higher vertebrates. Furthermore, Br KO MEFs activated the AMPK pathway due to the energetic perturbation caused by CIII deficiency (Supplementary Fig. 6i). Finally, the observed reductions in CIII protein levels led to significant reduction of all CIII activity as measured by spectrophotometric measurement of RC enzymatic activities.35 (Fig. 6g), while CI, CII, and CIV activities were not significantly affected. Altogether, our data demonstrate that the loss of Br in zebrafish causes overt and lethal mitochondrial deficiency due to the specific loss of CIII in the respiratory chain.

**Discussion**

Our findings in this report demonstrate that SEPs are preferentially enriched in the mitochondria and play programmatic functions in oxidative respiration. We speculate that this size bias might simply reflect the endosymbiotic origins of the mitochondrion. However, since the mito-SEPs including BR found in this study are not conserved in prokaryotes, an alternative explanation lies in the energetic barrier imposed by the double membranes of the mitochondrion against protein import. Simply put, it is energetically less costly to import small proteins than large proteins. For an organelle that is primarily responsible for maintaining energy balance, such a strategy would allow the mitochondrion to lower the cost of its own biogenesis during a time of increased ATP demand, allowing homeostasis to be reached in a shorter time frame.

The mechanism with which BR promotes CIII assembly and/or stability also calls for greater in-depth investigation. The assembly of complex III in mammals is poorly understood and largely inferred from studies performed in yeast.36 In zebrafish, loss of Br caused complete attenuation of the fully assembled 500-kDa CIII, with no evidence of immature pre-CIII assemblies seen in mutants of late assembly factors37–39, arguing that Br acts early in the biogenesis and known sequence of CIII assembly before dimerization of assembly intermediates37. Consistent with a possible direct role of Br as a CIII assembly factor in vivo, BN- and SDS-PAGE demonstrated partial co-migration of endogenous Br with CIII, in zebrafish mitochondria (Supplementary Fig. 6j, k). However, our data currently cannot exclude the possibility that BR mediates CIII biogenesis by facilitating the import of its constituent subunits or their stability following translation/import, which will require further investigation. BR is highly responsive to energy status. Electrons transferred from the oxidation of divergent fuel types converge at CIII. Because CIII is the center of the respiratory chain, we speculate that BR, acting as an early CIII assembly factor, coordinates the bioenergetic output of the respiratory chain with nutrient availability to maintain energy homeostasis. Conversely,
**Fig. 6 Br is required for complex III assembly and activity.**

**a** Quantitative proteomic analysis of WT and br zKO isolated adult skeletal muscle (SKM) mitochondria. Mean values of log₂ fold change (zKO/WT) for each detected protein from three biological replicates are plotted in a volcano plot. Subunits of each of the ETC complexes are labeled in the indicated colors. CI–CV = Complex I–Complex IV. *p*-values from two-sided unpaired t-test.

**b** The median log₂ fold change (zKO/WT) of all detected ETC proteins are further depicted in a dendrogram heatmap and separated by complex membership. AF = assembly factors.

**c** The median log₂ fold change (zKO/WT) of CI–CIV proteins were superimposed onto the structures of human CI–CIV to visually depict the observed changes. The same scalebar from panel 6b applies. Undetected proteins are shown in gray.

**d** SDS-PAGE of purified WT and zKO SKM mitochondria followed by immunoblotting for the indicated ETC proteins. Image is representative of two experiments.

**e** Quantitation of SDS-PAGE in 6d. Data represent mean and SEM. *n* = 5, except for Uqcr2b where *n* = 3. *p*-values from two-sided unpaired t-test.

**f** BN-PAGE and western blot analysis of WT and zKO SKM mitochondria. Position of free III₂ (red) and III₂+IVₙ (black) supercomplexes are indicated by arrowheads.

**g** Spectrophotometric respiratory chain activity assay (RCA) of WT and zKO SKM mitochondria. Data represent mean and SEM. Each dot represents one animal, *n* = 12 for WT and zKO animals. *p*-values from two-sided unpaired t-test.
BR levels affect cellular AMPK status. This raises the possibility that BR, through regulation of AMPK activation, might also be involved in regulating intra-hepatic lipid accumulation, a hypothesis that remains to be investigated. Recently, CIIH integrity has been implicated in regulatory T cell function and the prevention of autoimmunity, raising the intriguing possibility that BR can also coordinate nutrient availability with immune function. Lastly, congenital CIIH defects resulting from mutations in CIIH subunits or assembly factors cause very rare but severe mitochondrial disease characterized by growth retardation, lactic acidosis, aminoaciduria and early death.

The phenotypic similarities of the br KO zebrafish sets the basis to screen unresolved cases of mitochondrial disease for pathogenic variants in the BR (C12orf73) gene and potentially the genes encoding the other mito-SEPs described here.

In conclusion, by performing a screen of the SEP peptidome, we report the discovery of BRAWNIN, a mitochondrial peptide that is central to vertebrate oxidative phosphorylation. This emerging theme of the preponderance of SEPs in the mitochondria is seen also in the human heart, pointing to the in vivo significance of mito-SEPs in human physiology. When this manuscript was in revision, Saghatelian and colleagues reported the independent discovery of PIGBOS (MSU1 in this study) as a mitochondrial peptide that mediates the unfolded protein response in the endoplasmic reticulum, underscoring the biological diversity of mito-SEPs. Of high priority is deducing the contribution of the other 14 mito-SEPs to oxidative metabolism. Our work demonstrates the promise and value of mining and characterizing the sORF-encoded peptidome, which should motivate similar efforts to functionally characterize remaining uncharacterized SEPs, also known as Small GEms hidden in Plain view.

**Methods**

**sORF selection criteria.** sORF-encoded peptides (SEPs) were filtered using ribosome profiling-based metrics derived from sORFs.org, adapting the guidelines from Oleksiak et al. In all, 2343 human SEP genes were encoded by annotated short CDS genes, lincRNAs, upstream ORFs (uORFs), and intergenic ORFs that passed ribosome sequencing quality thresholds were selected for mitochondrial prediction (Ensembl annotation bundle v92 is applied). From the pool of 115,223 predicted sORFs, a minimum ORF score of 0.6 was imposed on nucleus encoded sORFs. The filtering criteria are set separately dependent on the sORF annotation. For the sORFs annotated in CDS genes, a minimum threshold of 50% in-frame coverage and a coverage uniformity between +1 and −1 was imposed, whereas for sORFs located on lincRNA, an in-frame coverage of 50% and a coverage uniformity between −0.5 and 0.5 was imposed. For sORFs located in the 5′-UTR region of protein coding genes, a minimum in-frame coverage of 0.7, a coverage uniformity between −0.3 and 0.3, a minimum length of 20 a.a. and a good FLOSS categorization was imposed. Also, these sORFs should be identified/predicted in at least 10 different Ribo-seq datasets. For intergenic sORFs (i.e. in between genes), a minimal up- and downstream gene distance of 1000 nucleotides (>1 kb), a minimal in-frame coverage of 0.9, a coverage uniformity between −0.3 and 0.3, a minimum length of 20 a.a., a good FLOSS categorization, and a minimum number of identifications in at least 1 Ribo-seq dataset was imposed. This resulted in a pool of 2343 sORFs encoded from 1027 different genes (Fig. 1C). See Supplementary text for more explanation of selection parameters.

**Mitochondrial prediction by WGCNA and GSEA.** RNA-seq datasets from three human tissues known to be enriched in mitochondrial activity were selected to enable mitochondrial functional prediction: non-neoplastic liver (21 samples from GSE94660 – RPKM normalized), normal heart left ventricular tissue (92 samples from GSE14670 – DESeq normalized counts). To compare those mitochondria-enriched tissues with a non-enriched one, a sun-protected skin dataset (91 samples from GSE85861) was selected for mitochondrial prediction (Ensembl annotation bundle v92 is applied). We performed WGCNA with the WGCNA package in R for all classifier and training set genes in all three datasets. i.e. we ranked all genes in a dataset according to their Spearman’s correlation with each test gene. Next, we performed GSEA using 7 gene set collections from MSig DB (Broad Institute)[23, 24], c2.cp.reactive, c2.geps, c2.epp.kegg, c2.ecc, c3bp, c5.mlf, and halluc, which gather 10,260 gene sets. Using GSEA package in R, the NES score of each gene set was calculated for each ranked list. NES was set to 0 if the GSEA output p-value (using an empirical phenotype-based permutation test procedure)[25] was not significant ($p > 0.05$). To decrease noise and increase specificity to mitochondrial activity, only gene sets with at least 100 genes with an average NES above 0.5 and gene sets containing more than 1000 genes were removed. Average, each gene had 63, 85, and 118 NES values in the heart, liver and SKM dataset respectively. These scores were used as variables for principal component analysis (PCA) using stats package. Altogether, each PCA analysis contained 3090 observations (1956 classifier+139 training +995 SEP candidate genes). Only 995 of the 1027 encoding 2343 SEP candidates were expressed in any one of the three datasets and could be tested by this method.

K-means clustering was applied to the PCA for a cluster number varying from 2 to 15. For each iteration, clusters were ranked from the highest percentage of mito genes to the lowest. To know if a cluster should be considered as positive or negative for mitochondrial identity, each cluster’s percentage was tested as a potential threshold. The threshold which led to the highest number of correct prediction of the training set was kept. Clusters with a mitochondrial percentage below the threshold were considered negative, otherwise positive. Candidates were given a binary score of 0 or 1 depending on whether they belonged to a negative or positive cluster respectively. This process of binary scoring was made for every cluster and later, each SEP candidate was given an average k-means score was set to −1 if the gene was not encoded in the dataset. Based on the k-means scores of the training set, it appeared that 75% was the minimum average k-means score needed for a peptide to be predicted as mitochondrial. To be more stringent on this prediction we added the criteria of confidence level. “High confidence” meant that a SEP candidate required an average k-means score above 75% of all RNAseq where it is expressed. Refer to Supplemental Text for more information.

**Mitochondrial prediction by targeting motif prediction.** Mitochondrial targeting motifs were predicted in peptides that were not selected by WGCNA/GSEA. Different types of mitochondrial targeting motifs were predicted using separate algorithms. Classical mitochondrial targeting sequences (MTSs) were captured by their homology prediction. TargetP (mitochondria, Reliability score > 0.5) and Mitofate (score > 0.8) or Mitoprot (score > 0.8) were considered as positive predictions. Mitochondria targeting signal, but they are poorly predicted by algorithms used above because these sequences are not cleavable. These hydrophobic α-helices were predicted by TMHMM or SignalP (1.1 and 5)24,25. To eliminate secretory proteins, the maximum TMD hydrophobicity was set at 0.6, as we noticed that mitochondrial TMDs are on average less hydrophobic compared to non-mitochondrial TMDs. In addition, mitochondrial intermembrane proteins often contain twin cysteine pairs (two CXXC, two CXCX, or one CXCX and one C10X0, where Δ marks any residues in between and apical cysteine residues)26, and these internal signals are generally missed by prediction algorithms. We thus included peptides with these characteristic cysteine pairs in the screen.

**Mitochondrial prediction by mitochondrial peptide RESPIN.** Quantitative mass spectrometry data on isolated mitochondria are available at the PRIDE public proteomics data repository (accession number PXD004660) and downloaded using the PRISE API25. Subsequently, the raw data files were converted and filtered using MConvert24. Next, the SearchGUI tool was used to perform the peptide to spectrum matching, based on a combination of two database search algorithms X!Tandem and MS/MS+1 against a custom sequence database. This database consists of (i) the UniProt human functional proteome including all isoforms27, (ii) the human sORFs from the sORFs.org database28, and (iii) the clash database (https://www.the-cdb.org). SearchGUI queries were adapted from the original study29. Matched PSMs were validated using PeptideShaker30, only PSM solely mapping to sORF peptide sequences with a spectrum coverage of at least 30% and FDR<0.01 were retained. After selecting peptides by the three strategies, redundant isoforms were manually removed if they shared identical sequences with the lowest score identified. The consensus prediction of TargetP (mitochondria, Reliability class 1 & 2) and Mitofate (score > 0.8) or Mitoprot (score > 0.8) were considered as positive predictions. Mitochondria targeting signal, but they are poorly predicted by algorithms used above because these sequences are not cleavable. These hydrophobic α-helices were predicted by TMHMM or SignalP (1.1 and 5)24,25. To eliminate secretory proteins, the maximum TMD hydrophobicity was set at 0.6.

**SEP transient over-expression.** The open reading frames (ORFs) of selected candidates were synthesized and inserted into the NheI-XhoI site of a pcDNA3.1 (+) vector, under the control of the CMV promoter. Peptides were tagged with a single HA tag distal to their predicted targeting motifs or c-terminus if no motif was predicted. ORFs with non-AUG start codons were replaced with a AUG start codon to avoid including cognate 3′UTRs required for recognizing near-cognate start sites31. HEK293 cells were seeded in 24-well plates and transfected with FuGENE (Promega) and stained as described in the immunofluorescence section.
Zebrafish Br knockout generation. *zBrawnins* are encoded by the sickness21-16aa17 gene which contains 3 exons with the ORF in exons 2 and 3. EST mining on NCBI and Ensembl mapping in this region in the zebrafish genome did not reveal the presence of a duplicated allele or paralogue. To completely excise the ORF, we used ZfET52 to design two gRNAs 5’ and 3’ to the ORF with recognition sites underlined. ATG in bold marks the translation start site.

5’ - Br gRNA: GAAATATACGACTCACTATAATGAAAATAGCACGTTTTT

3’ - Br gRNA: AAAATGGAATATACGACTCACTATAATGAAAATAGCACGTTTTT

WT undeleted bands yield a 624 bp band while the deleted allele (ΔBr) yields a 62 bp band with HotStart Taq Polymerase (Qiagen). In all, 2 Br founders (ΔBig and ΔSmall) with complete deletions were brought forward for characterization. Both founders showed similar growth defects. All data presented are derived from ΔBig but are representative of ΔSmall, which had a more severe phenotype.

Zebrafish husbandry and growth tracking. Zebrafishes were reared under standard conditions, 28°C with approved protocols following regulations stipulated by the IACUC committee of A*STAR Singapore. Heterozygous crosses were performed starting at the F2 generation to produce WT, Hets, and zygotic KO larvae. These were housed at a density of 50 per tank. In all, 15–20 animals were randomly selected and genotyped every week by scale to track Mendelian representation over time under conditions of unsegregated growth. Briefly, scales were lysed in 50 mM NaOH at 95°C for 15 min followed by neutralization with 1/10th volume Tris-HCl pH 8.0 and debris pelleted by centrifugation at 3000 × g for 5 min before PCR with LHW4_zfBr_5_F and LHW7_zfBr_3_R primers listed above. To enable segregative growth, larvae were genotyped at 3dpf picking the tip of the tailfin3 and housed at a lower density of 20/tank. KO fish were given one extra feed per day to boost their growth. The standard length of these larvae were measured once every week under a stereoscope (up to day 28) or with a ruler after 3 weeks of age, to provide a robust representation of their growth due to their low body weight and lethargic phenotype. Female Br KO females are however, still able to mate with WT or Het fish, demonstrating preservation of their fertility.

Birefringence analysis. In all, 4–6 dpf larvae were anaesthetized briefly with tricaine before being placed in 30% methanol in egg water. Imaging was performed by placing the dish containing the larvae in between a pair of polarized filters (50 mm Diameter Unmounted, Linear Glass Polarizing Filter, Stock No. #43787 from Edmund Opticals) aligned at 90° to each other. Images were collected under brightfield with a Nikon SMZ800 stereo connected to a CCD camera.

Zebrafish motility monitoring. In all, 5 dpf mz KO and WT larvae were gently pipetted into a 48-well plate, each fish per well for monitoring locomotor activity. The 48-well plate was placed in a sound-attenuated incubator to isolate it from extraneous noise and light. Inside the incubator, a white-light LED box was positioned below the plate and videos were recorded from the top at a speed of 5.3 frames per second (3.3 fps) by a Basler USB camera in a resolution of 2048 × 1024 pixels. The position (x-y coordinates) of the fish was determined in real-time based on background subtraction algorithms written in Python utilizing OpenCV library. This data was then exported into Microsoft Excel files and analyzed offline for speed and distance measures.

Zebrafish respiratory analysis. Manually dechorionated 24 hpf (1 dpf) mz KO larvae were placed into each well (1 well) of an Agilent XF96e Spheroid microplate containing 180 µl of sea salt water and allowed to adapt for 1 h. Basal respiration was recorded on the Seahorse XF96e at 28°C. This was achieved by placing the XF96e in a refrigerated chamber held at 16°C. We were unfortunately unable to perform a Mito Stress test on live larvae because the repeated mixing steps following injection of Oligomycin and FCCP caused significant lethality to the larvae due to the low height clearance of the XF96e oxygen probe.

Zebrafish metabolomic analysis. At 5 dpf, 50 WT or mz KO larvae were collected into a tube under normal conditions, placed on ice for 5 min to immobilize larvae, followed by gentle sedimentation. All egg water was removed and larvae were rinsed washed twice with ice cold water. The dried pellet was weighted to 2 decimal points in mg and snap frozen, thawed on ice and homogenized in 50% H2O/50% Acetonitrile using zirconia beads in a liquid nitrogen cooled rotor (2 cycles of 3000 × g for 20 s). For organic acid extraction, 300 µl of tissue homogenate was extracted with ethylacetate, dried and derivatized with N, O-Bis(trimethylsilyl)trifluoroaceticamide, with protection of the alpha keto groups using trimethylsilylamine (Sigma-Aldrich). The mixture was quilibrated for 30 s and 1.2 ml of HPLC grade methanol was added to the mixture. After vortexing, the mixture was incubated in 50°C for 10 min and centrifuged to pellet the precipitated protein. The supernatant was removed and dried under nitrogen in a clean microcentrifuge tube. The dried extract was reconstituted in 50 µl of methanol before use. Using a liquid chromatography-triple quadrupole mass spectrometer. Trimethylsilyl derivatives of organic acids were separated by gas chromatography on an Agilent Technologies HP 7890A and quantified by single ion monitoring on a 5975C mass spectrometer using stable isotope dilution. The initial GC oven temperature was set at 70°C, and ramped to 300°C at a rate of 40°C/min, and held for 2 min.

MitoCore modeling. The MitoCore model was downloaded in SBML format from Zieliński et al.32. Each of the 485 reactions in the model was blocked (setting upper and lower bounds of permissible flux to zero), one reaction at a time, and the ATP production (Reaction ID: "OE_ATP_MitoCore") was constrained to have a flux >90% of the theoretical capacity of the blocked model. Uniform sampling was performed for each blocked model using the CHRR algorithm of Cobra toolbox 3.0 to obtain 5000 flux vectors that represent the solution space. As baseline comparison, the unchanged MitoCore model (unconstrained model) was also sampled to obtain 5000 flux vectors. The flux distributions for the export reactions of Lactate (Reaction ID: "L_LACt2r") and Succinate (Reaction ID: "SUMt_MitoCore") were compared between the unconstrained model and each of the 485 constrained models using the Wilcoxon ranksum test, and the effect sizes were computed using Cohen’s d. Reaction blockages that produced a negative effect size for either lactate or succinate export were discarded, and the remaining blocked models were ranked by the sum of effect sizes (for Lactate and Succinate). A similar workflow was repeated to identify which of the 371 genes in the MitoCore model had largest effect size to increase succinate and lactate export when knocked out (Single Gene Deletion studies using Cobra Toolbox 3.0)34.

Mitochondria flux assays. Respirometric analysis of purified mitochondria from zebrafish skeletal muscle was performed using Agilent’s Seahorse platform as described by Boutagot et al.35. Briefly, 1.0–1.4 µg of purified mitochondria were plated in each well of a XF96e cell culture plate. To measure uncoupled electron flow, assay buffer contained 5 mM pyruvate, 4 mM carbonyl cyanide 4-(trifluoromethoxy) phenylhydrazone (FCCP), 1 mM malate. Rotenone (2 µM, final), succinate (5 mM, final), antimycin A (4 µM, final), TMPS/acascorbic acid (100/100 µM) were injected sequentially during the assay. In coupling assay with pyruvate and malate, assay buffer contained 10 mM pyruvate and 5 mM malate. In coupling assay with Complex 1 inhibited, assay buffer contained 10 mM succinate and 2 µM rotenone. In coupling assay, ADP, oligomycin A, FCCP and antimycin A were injected into assay plate sequentially to 3.2 mM, 3 µM, 4 µM and 4 µM, respectively. Citrate synthase activity (described below) was used for post-normalization of OCR rates.

Mitochondria purification from cultured cells and tissues. Mitochondria were isolated from cultured cells using mitochondria isolation kit (Abcam, ab101168). Cells were harvested and ruptured with a Dounce homogenizer in the buffer A containing Complete protein inhibitor cocktail (Roche) and 1 mM phenylmethylsulfonyl fluoride (PMF). 15 strokes were performed with a motorized pestle operated at 300 rpm. Nuclei and unbroken cells were removed by centrifugation at 600 × g. Homogenates were spun at 7000 × g for 15 min to precipitate mitochondria. Mitochondria pellets were collected for downstream uses.
To isolate mitochondria from zebrafish, skeletal muscles of fish were dissected and minced. Tissue was disrupted in muscle mitochondria isolation buffer (67 mM sucrose, 50 mM KCl, 1mM EDTA, 0.2% fatty acid free BSA, 50 mM Tri-HCl, pH 7.4) with a Dounce homogenizer tight pestle operated at 1300 rpm. The extract was centrifuged at 600 x g for 5 min to clear intact myofibrils and heavy cell debris and then spun at 7000 x g for 15 min to isolate the desired mitochondrial fraction.

**Antibody generation.** Custom polyclonal antibodies were generated against human BR (hBR) and zebrafish BR (zBr) using the following peptide antigens. For SDS-PAGE, BN-PAGE, IF and IHC of human/mouse BR: 5′-EVHYRTPDPITPEIPPKGELKTEC-3′ for SDS-PAGE of zBr; 5′-RPDLSIPPEIPPKGELKTEC-3′ and secondary antibodies was prepared in 3% BSA. IRDye 800CW goat anti-rabbit and anti-mouse IgG at 1:5000. Immunofluorescence and immunohistochemistry. Cells were fixed by 4% PFA at 37°C for 20 min and washed with PBS twice. Where appropriate, cells were permeabilized with 25 mM of Mitotracker Red at 37°C for 30 min before the fixation. Samples were solubilized with 0.3% Triton X at room temperature for 10 min and washed twice with PBS. Cells were blocked with blocking buffer (3% BSA, 10% PBS in PBS) at room temperature for 1 h. Primary antibodies were diluted to desired concentration in the blocking buffer and incubated with cells for overnight, rolling at 4°C. Proteins of interest were probed with the following antibodies: anti-BR, 1:200 (Novus, NBP1-90536); anti-ATP5A (AbClonal, A5349); anti-CS, 1:2000 (Santa Cruz, SC-390693); anti-TOMM20, 1:5000 (Pro-Blott). Membranes were scanned using Licor Image Studio.

**Immunofluorescence and immunohistochemistry.** Cells were fixed by 4% PFA at 37°C for 20 min and washed with PBS twice. Where appropriate, cells were permeabilized with 25 mM of Mitotracker Red at 37°C for 30 min before the fixation. Samples were solubilized with 0.3% Triton X at room temperature for 10 min and washed twice with PBS. Cells were blocked with blocking buffer (3% BSA, 10% PBS in PBS) at room temperature for 1 h. Primary antibodies were diluted to desired concentration in the blocking buffer and incubated with cells for overnight, rolling at 4°C. Proteins of interest were probed with the following antibodies: anti-BR, 1:200 (Novus, NBP1-90536); anti-ATP5A (AbClonal, A5349); anti-CS, 1:2000 (Santa Cruz, SC-390693); anti-TOMM20, 1:5000 (Pro-Blott). Membranes were scanned using Licor Image Studio.

**Sodium carbonate extraction.** Sodium carbonate extraction was performed as described with minor modifications. Mitochondria pellets isolated from HEK293T cells were resuspended in buffer C (320mM Sucrose, 1 mM EDTA, 10 mM Tris-Ci pH 7.4) and incubated containing 1% Triton X or 0.1M NaClO4 pH 11. The suspensions were incubated at 4°C for 1 h with gentle rotation. Supernatants and pellets were separated by centrifugation at 20,000 × g for 1 h at 4°C. The supernatant was gently removed without disturbing the pellet. Equal fractions of supernatant and pellet were analyzed by SDS-PAGE immunoblotting.

**Mitochondrial membrane fractionation.** Mitochondria were isolated with isoionic buffer (250 mM sucrose, 1 mM EDTA, 10 mM HEPEs KOH pH 7.4, protease inhibitor cocktail and 1 mM PMSF) from HEK293T cells following procedures described in the mitochondrial isolation section. Mitochondria pellet was gently resuspended in the isoionic buffer to get protein concentration at 1 mg/mL. Aliquots of mitochondrial were solubilized by increasing concentrations of digitonin (0%, 0.1%, 0.115%, 0.13%, 0.145%, 0.16%, 0.175%, 0.19%, 0.205%, and 0.22%) for 1 h at 4°C. Soluble and insoluble fractions were separated by centrifugation at 20,000 × g for 20 min Insoluble fractions were fully resuspended in an equal amount of isoionic buffer to match the volume of supernatant. Both fractions were lysed in 1X laemmli sample buffer and analyzed by SDS-PAGE and immunoblotting.

**Protease sensitivity assay.** Mitochondria were isolated with isoionic buffer (250 mM sucrose, 1 mM EDTA, 10 mM HEPEs-KOH pH 7.4, protease inhibitor cocktail) from HEK293T cells. Mitochondria pellet was washed to eliminate protease inhibitors. Mitochondria pellets were resuspended to a protein concentration of 1 mg/mL. Mitochondrial membranes were by differentially solubilized by increasing concentrations of digitonin (0–0.2% as indicated in the figure) for 10 min on ice. Mitochondria were added to 100 μg/mL and incubated for 30 min on ice to allow complete digestion of accessible proteins. To terminate the protease digestion PMSF was freshly prepared and added to a concentration of 8 mM. Samples were analyzed by western.

**Mass spectrometry analysis of isolated mitochondria.** Crude mitochondrial protein isolated from adult zebrafish skeletal muscle or MEF was normalized using the Pierce Protein Assay Kit (ThermoFisher Scientific). Mitochondria were solubilized in 1% (v/v) SDS, 100mM Tris pH 8.1, 40 mM chloroacetamide (Sigma) using 1× Laemmli buffer (ThermoFisher Scientific). Mitochondria were centrifuged at 50,000 × g for 5 min at 99°C with 1500 rpm shaking followed by 15 min sonication in a water bath sonicator. Protein digestion was performed using trypsin (ThermoFisher Scientific) at a 1:50 trypsin:protein ratio at 37°C overnight. The supernatant was transferred to StageTips made in-house containing 3 μL trypsin plugs of 3M "Empore" SB-RPS substrate (Sigma) as described previously with modifications. Briefly, isopropanol (99% v/v) containing 1% (v/v) TFA was added to the tip and mixed with the lysate prior to centrifugation at 3000 × g. StageTips were then washed first with the same solution followed by an additional wash with 0.2% (v/v) TFA. Peptides were eluted in 80% (v/v) acetonitrile (ACN) and 1% (v/v) NH4OH, and then exchanged to a final concentration of 1% TFA prior to drying in a CentriVap Benchtop Vacuum Concentrator (Labconco). Peptides reconstituted in 0.1% TFA were desalted on Pierce C-18 tips (Thermo Scientific) as per manufacturer’s instructions prior to drying under vacuum. Desalted peptides were reconstituted in 100mM triethylammonium bicarbonate (TEAB, Thermo Fisher Scientific) and tandem mass spectra acquired in 8:1 label:protein ratio as per manufacturer instructions. Labels were pooled such that each multiplex set contained three replicates each of Br KO and
control mitochondria. Pooled samples were fractionated using the Pierce High pH Reversed-Phase Peptide Fractionation Kit (ThermoFisher Scientific) as per manufacturer’s instructions for a TMT experiment. Individual fractions were dried under vacuum and reconstituted in 2% (v/v) acetonitrile (ACN) and 0.1% (v/v) trifluoroacetic acid (TFA).

Liquid chromatography (LC) coupled MS/MS was carried out on an Orbitrap Lumos mass spectrometer (ThermoFisher Scientific) with a nanoESI interface and a Dionex nanoLC system. The LC system was equipped with an Acclaim Pepmap nano-trap column (Dionex-C18, 100, 75 μm × 2 cm) and an Acclaim Pepmap RSLC analytical column (Dionex-C18, 100, 75 μm × 30 cm). The pump flow was set to 0.3 μL/min, protein lysate stored in 1% (v/v) ACN and 0.1% (v/v) formic acid (solvent A) and 5% DMSO in 100% v/v ACN and 0.1% (v/v) formic acid (solvent B) using a custom 128-minute non-linear gradient that included a step for equilibration of the column prior to injection of the next sample. The Synchronous Precursor Selection (SPS)–MS3 based TMT method was used. In brief, a full MS1 spectra was acquired in positive mode at 120,000 resolution scanning from 380–1500 m/z. AGC target was at 4e5 and maximum injection time of 50 ms. Precursors for MS2/MS3 selection were based on a Top 3 second method. MS2 analysis consists of collision induced dissociation (CID) source isolation waveforms frequency notches. The MS3 precursors were fragmented by collision induced dissociation (CID) in the ion trap. Following acquisition of each MS2 spectrum, multiple MS2 fragment ions were captured in the MS3 precursor population using isolation waveforms frequency notches. The MS3 precursors were fragmented by high-energy collision-induced dissociation (HCD) and analyzed using the Orbitrap with source isolation waveforms frequency notches. The MS3/MS2 isolation waveforms were adjusted to capture unique peptides. Zebrafish mitochondria was processed in Perseus (version 1.6.10.43) and searched against UniProt zebrafish database82.

Mitochondria were isolated from tissues as described above and solubilized with IP buffer (150 mM NaCl, 1 mM EDTA, 50 mM HEPEs-KOH pH 7.4) containing 1% N-Dodecyl β-α-maltoside (DDM), cOmplete, EDTA-free protease inhibitor cocktail (Roche) and 1 mM PMSF at 4 °C overnight at 37 °C with shaking. The sample was hydrolyzed by adding TFA to 1% and then by adding 2.5 μg of Tryptsin (Promega) for 3 h at 37 °C with shaking and then treated with 2.5 μg of Trypsin (Promega) for over-night at 37 °C with shaking. The sample was then desalted on a C18 cartridge (Oasis) and eluted with 50% ACN and 100 mM TEAB pH 8.5, vacuum concentrated and resuspended in 100 mM TEAB pH 8.5. Peptides were labeled with 10 plex tandem mass tags (TMT) for over-night and then quenched by 0.5% Tri-HCl pH 7.4, pooled, desalted again and analyzed by an Orbitrap LQ/LQ50.

Mouse housing conditions. In this study, C57BL6/J mice were housed on a 12-h light/dark cycle with water and normal chow diet provided ad libitum. Room temperature was maintained at 21–24 °C and relative humidity at 40––60%. All animal procedures were approved and conducted in accordance with the Duke-NUS,SingHealth Institutional Animal Care and Use Committee.

BR co-IP with complex III. To express TFL tagged BR from mouse tissue, BR-associated virus (AAV) expression construct was generated by inserting the murine BR ORF into a pcscAAV backbone with a single TFL peptide fused to the C-terminus of BR. AAV9 pseudotyped virus was produced by Vector Core, Genome Institute of Singapore, Singapore. In all, 105 viral genomes were constituted in 10 μL PBS and administrated to the gastrocnemius muscle of 1-month-old female animals. Skeletal muscles were isolated 1–2 months after the AAV administration. Mitochondria were isolated from tissues as described above and solubilized with 1% digitonin (mitochondria protein ratio was normalized to 4 g/g) in IP buffer (150 mM NaCl, 1 mM EDTA, 50 mM HEPEs-KOH pH 7.4) containing cOmplete, EDTA-free protease inhibitor cocktail (Roche) and 1 mM PMSF at 4 °C for 30 min with constant rocking. The insoluble fraction was removed by centrifugation at 20,000 × g for 15 min. The supernatant was incubated with 30 μl of anti-FLAG M2 resin and rotated for over-night in a cold room. The resin was extensively washed with IP buffer supplemented with 0.1% digitonin for four times. BR-FLAG and its interaction partners were eluted with 0.5 mg/mL 3XFLAG peptide (Sigma) in IP buffer at 10 min with elute and eluate were analyzed by SDS-PAGE. In the reciprocal IP, mitochondria were isolated from HEK293T cells transiently expressing UQCRQ-FLAG and BR-HA. Protein IP was described as above.

Cell culture. All cell lines were obtained from American Type Cell Collection (ATCC; www.atcc.org) and cultured using standard tissue culture techniques. U87MG (ATCC® CRL-1330) (HyClone) and 1% of 10,000 U/mL of Penicillin-Streptomycin (HyClone). U87MG (ATCC® HTB-14) (HyClone) were cultured in Minimal Essential Medium with L-Glutamine ( Gibco). All media were supplemented with 10% American South American sourced Fetal Bovine Serum (HyClone) and 1% of 10,000 U/mL of Penicillin-Streptomycin (Gibco). Starvation regimes were performed as follows: HEK293T cells were plated in normal DMEM (4500 mg/L glucose) overnight, followed by a complete media change to that containing 5 mM Temozolomide in complete media, lacking FBS (Serum- ) and or DMEM without glucose (GLUCOSE-).
passed through a 70-μm cell strainer (BD). Cells were seeded to a 10-cm dish pre-coated with gelatine and was cultured with complete medium. In sub-culturing, MEFs were detached with 0.05% Trypsin-EDTA (Gibco).BR over-expression by lentivirus transduction. The human BR ORF was inserted into a pCDH lentivirus vector under the control of a CMV promoter. Plasmid was then packaged into viral particles using the pPACKH1 lentivector Packaging Kit (System BioScience) in HEK293T cells. In all, 25% of total crude virus extracted was used to infect U87MG for 2 days before G418 (Santa Cruz) selection at 12 μg/mL.

Brawnin knockdown. siRNA-meditated knockdown of BR was performed in U87MG with ON-TARGET plus Human C12orf73 (728568) siRNAs (Dharmacon) with the following sequences: 5′-CACUGCAUGUUAUCGGAU-3′ (siBR_14) and 5′-GAGCGGCAGCUCCAGAACU-3′ (siBR_16). Each equivat ratio of both targeting siRNAs are mixed to achieve maximal knockdown. Control siRNA sequence is 5′-UGUUUCAUAGGUGCUAA-3′. Briefly, siRNAs are transfected at 25 nM final concentration using Dharmafect using a suspension transfection protocol for U87MG. siRNA-mediated knockdown or BR was performed using the following sequences: 5′-AAAATGTCCGGGTATAGCCTTCTAGT-3′ (shBR_1) and 5′-AAAAATGTGATTTAGCCTGAACT-3′ (shBR_3). siRNA hairpins were cloned into a lentiviral backbone to allow stable expression. Viral particles were produced in LentiX293 and purified according to standard procedures. 500,000 U87MG cells were collected and resuspended with respective medium with 32 μg/mL of polybrene, and added to polybrene-treated cells at a target MOI of 3. Cell virus suspension solution was centrifuged at 1000 × g for 1 h at room temperature to maximize the collision between cells and viruses. The mixture was then diluted four times with respective medium and seeded to one 6-well plate (Costar). Following recovery from viral transduction, cells were selected with 1 μg/mL of puromycin until a non-transduced control showed complete death. U87MGScr shBR cells were used within 10 passages of initial transduction to minimize phenotypic drift due to compensation.

Seahorse cultured cell MitoStress test. In all, 15,000 U87MG cells were plated on poly-1-Lysine (Sigma) treated Seahorse XP96 Cell Culture Microplates (Agilent) and XF96 sensor cartridge (Agilent) was hydrated with distilled water at 37°C with no CO2, the day prior to MitoStress Assay. On the day of actual MitoStress Assay, cells were washed once and media were replaced with XF basal DMEM supplemented with 1 mM pyruvate (Sigma), 2 mM glutamine (Sigma), and 10 mM glucose (Suprema). Pre-hydrated sensor cartridge was calibrated with Seahorse Wave software, Agilent. Citrate Synthese Normalization Assay was performed on the cultured cell plate to normalize the measured Oxygen Consumption Rate (OCR) and Extracellular Acidification Rate (ECAR).

Citrone Synthese Normalization Assay. Citrate Synthese Normalization Assay was performed on Seahorse XP96 Cell Culture Microplate (Agilent) right after seahorse assay or on plates frozen at –80°C with leftover medium aspirated. All leftover medium was removed and replaced with 113 μL of CS buffer (200 mM Tris buffer at pH8.0, 0.2% Triton X-100 (v/v), 100 μM DNB (Sigma) in 100 mM Tris buffer at pH8.0, 1mM Acetyl-CoA (Sigma)) per well. In all, 5 μL of 100 μM Oxa-lOacetate (Sigma) in water was added to each well as reaction substrate. Absorbance at 412 nm at 37°C was recorded with Tecan Microplate Reader M200 (Tecan) at minimal time interval for 8 min Citrate Synthese activity was then calculated using the formula below

\[ \text{CS activity} = \frac{n}{\sum_{i=1}^{n} \frac{(A_i - A_f)}{t_i - t_f}} \cdot \frac{n}{n - 1} \]

\( n \) is the total number of absorbance records; \( A \) is the absorbance recorded; \( t \) is the time.

Respiratory chain enzymatic assays. Respiratory chain enzymatic assays were performed with zebrafish skeletal muscle homogenates according to the protocol as described by Spinazzi et al.13.

Statistics and reproducibility. Unless otherwise stated, all data are presented as Mean ± standard error of mean (SEM). Statistical approaches to test differences of means were performed using Prism 5.0. The statistical test used in each panel is indicated in each accompanying figure legend. Where p-values are not explicitly highlighted, significance levels follow the convention of *p < 0.05, **p < 0.01, ***p < 0.001.

Received: 19 July 2019; Accepted: 14 February 2020;
Published online: 11 March 2020.

References
1. Ingolia, N. T., Ghaemmaghami, S., Newman, J. R. S. & Weissman, J. S. Genome-wide analysis in vivo of translation with nucleotide resolution using ribosome profiling. Science 324, 218–223 (2009).
2. Ingolia, N. T. et al. Ribosome profiling reveals pervasive translation outside of annotated protein-coding genes. Cell Rep. 8, 1565–1579 (2014).
3. Buzzini, A. A. et al. Identification of small ORFs in vertebrates using ribosome footprinting and evolutionary conservation. EMBO J. 33, 981–993 (2014).
4. Couso, J.-P. & Pataquim, P. Classification and function of small open reading frames. Nat. Rev. Mol. Cell Biol. 18, 575–587 (2017).
5. Ruiz-Orera, J., Meseguer, X., Subirana, J. A. & Alba, M. M. Long non-coding RNAs as a source of new peptides. Elife 3, e05252 (2014).
6. Hao, Y. et al. SmPro: a database of small proteins encoded by annotated coding and non-coding RNA loci. Brief Bioinform. 19, 636–643 (2018).
7. Olexiouk, V. et al. sORFs.org: a repository of small ORFs identified by ribosome profiling. Nucleic Acids Res. 44, D324–D329 (2016).
8. van Heesch, S. et al. The translational landscape of the human heart. Cell 178, 242–260.e29 (2019).
9. Czeglédy, S., Ho, L., Tian, J. & Reversade, B. ELABELA: a hormone essential for heart development signals via the apelin receptor. Dev. Cell 27, 672–680 (2013).
10. Pan, W. et al. CSFβ1/Conor99, a novel potential cytokine, inhibits colon cancer cell growth through inducing G1 arrest. Sci. Rep. 4, 6812 (2015).
11. D’lima, N. G. et al. A human microprotein that interacts with the mRNA decapping complex. Nat. Chem. Biol. 13, 174–180 (2017).
12. Bi, P. et al. Control of mRNA translation by the fusogenic microprotein myosinX. Science. https://doi.org/10.1126/science.aam9361 (2017).
13. Nelson, B. R. et al. A peptide encoded by a transcript annotated as long noncoding RNA enhances SERCA activity in muscle. Science 351, 271–275 (2016).
14. Herberg, S., Gert, K. R., Schleifer, A. & Pauli, A. The Ly6/uPAR protein function is necessary and sufficient for specific follicle maturation. Development 361, 1029–1033 (2018).
15. Ho, L. et al. ELABELA deficiency promotes preeclampsia and cardiovascular malformations in mice. Science 357, 707–717 (2017).
16. Calvo, S. E., Clausen, K. R. & Mootha, V. K. MitoCarta2.0: an updated inventory of mammalian mitochondrial proteins. Nucleic Acids Res. 44, D1251–D1257 (2016).
17. Akimoto, C. et al. Translational repression of the McKusick-Kaufman syndrome transcript by upstream open reading frames encoding mitochondrial proteins with alternative polyadenylation sites. Biochim. Biophys. Acta 1830, 2728–2738 (2013).
18. Stein, C. S. et al. Mitoregulin: an IncRNA-encoded microprotein that supports mitochondrial supercomplexes and respiratory efficiency. Cell Rep. 23, 3710–3720.e8 (2018).
19. Makarewich, C. A. et al. MOXI is a mitochondrial micropeptide that enhances fatty acid β-oxidation. Cell Rep. 23, 3701–3709 (2018).
20. Rathore, A. et al. MIEF1 microprotein regulates mitochondrial translation. Biochimie 57, 539–5575 (2013).
21. Dinyer, M. G. et al. The mitochondrial acyl-carrier protein interaction network highlights important roles for LYRM family members in complex I and mitoassembly assembly. Mol. Cell. Proteomics https://doi.org/10.1074.mcp.RA119.017849 (2019).
22. Nunnari, J. & Suomalainen, A. Mitochondria: in sickness and in health. Cell 149, 1145–1159 (2012).
23. Pham, A. H., McAffer, J. M. & Chan, D. C. Mouse lines with photoactivatable mitochondria to study mitochondrial dynamics. Genesis 50, 833–843 (2012).
24. Becker, T., Böttinger, L. & Pfanner, N. Mitochondrial protein import: from transport pathways to an integrated network. Trends Biochem. Sci. 37, 85–92 (2012).
25. Petersen, T. N., Brunak, S., von Heijne, G. & Nielsen, H. SignalP 4.0: discriminating signal peptides from transmembrane regions. Nat. Methods 8, 785–786 (2011).
26. Langfelder, P. & Horvath, S. WGCNA: an R package for weighted correlation network analysis. BMC Bioinformatics 9, 539 (2008).
27. Hertzberg, S. & Shaw, R. J. AMPK: guardian of metabolism and mitochondrial homeostasis. Nat. Rev. Mol. Cell Biol. 19, 121–135 (2018).
28. Corton, J. M., Gillespie, J. G., Hawley, S. A. & Hardie, D. G. 5-aminoimidazole-4-carboxamide ribonucleoside: A specific method for activating AMP-activated protein kinase in intact cells. Eur. J. Biochem. 229, 541–546 (1995).
29. Ha, J., Daniel, S., Royles, S. S. & Kim, K. H. Critical phosphorylation sites for acetyl-CoA carboxylase activity. J. Biol. Chem. 269, 22162–22168 (1994).
30. Calvo, S. et al. Systematic identification of human mitochondrial disease genes through integrative genomics. Nat. Genet. 38, 576–582 (2006).
31. Chinmery, P. F. Mitochondrial Disorders Overview. GeneReviews (1993).
32. Zazulaki, L. P., Morioka, A. C., Smith, R. A. & Robinson, J. A. Mitochondrial flexibility of mitochondrial respiratory chain disorders predicted by computer modelling. Mitochondrion 31, 45–55 (2016).
33. Tzagoloff, A., Wu, M. A. & Crivellone, M. Assembly of the mitochondrial membrane. Characterization of COR1, the structural gene for the 44-kilodalton core protein of yeast coenzyme QH2-cytochrome c reductase. J. Biol. Chem. 261, 17165–17169 (1986).
34. Oka, T. et al. Identification of a novel protein MICS1 that is Involved in maintenance of mitochondrial morphology and apoptotic release of cytochrome c. Mol. Biol. Cell 19, 2597–2608 (2008).
35. Spinazzi, M., Casarin, A., Pertegato, V., Salvati, L. & Angelini, C. Assessment of mitochondrial respiratory chain enzymatic activities on tissues and cultured cells. Nat. Protoc. 7, 1235–1246 (2012).
36. Signes, A. & Fernandez-Vizarra, E. Assembly of mammalian oxidative phosphorylation complexes I-V and supercomplexes. Essays Biochem. 62, 255–270 (2018).
37. Zara, V., Conte, L. & Trumpower, B. L. Biogenesis of the yeast cytochrome bc1 complex. Biochim. Biophys. Acta 1802, 2410–2418 (2010).
38. Zou, J., Hou, J., Hu, T. & Jiang, J. Single-molecule observation of human mitochondrial complex I. Nucleic Acids Res. 37, 1007–1022 (2009).
39. Strood, D. A. et al. Accessory subunits are integral for assembly and function of human mitochondrial complex I. Nature 538, 123–126 (2016).
40. Vizcaino, J. A. et al. 2016 update of the PRIDE database and its related tools. Nucleic Acids Res. 44, D447–D456 (2016).
41. Ternet, T. et al. How to submit MS proteomics data to ProteomeXchange via the PRIDE database. Proteomics 14, 2233–2241 (2014).
42. Chambers, M. C. et al. A cross-platform toolkit for mass spectrometry and proteomics. Nat. Biotechnol. 30, 918–920 (2012).
43. Barsnes, H. & Vaulou, M. SearchGUI: a highly adaptable common interface for proteomics search and de novo engines. J. Proteome Res. 17, 2552–2555 (2018).
44. Craig, R. & Beavis, R. C. TANDEM: matching proteins with tandem mass spectra. Bioinformatics 20, 1466–1467 (2004).
45. Kim, S. & Pevzner, P. A. MS-GF++ makes progress towards a universal database search tool for proteomics. Nat. Commun. 5, 5277 (2014).
46. UniProt Consortium UniProt: a worldwide hub of protein knowledge. Nucleic Acids Res. 47, D506–D514 (2019).
47. Olexiouk, V., Van Criekinge, W. & Menschera, G. An update on sORFs: a large repository of small ORFs identified by ribosome profiling. Nucleic Acids Res. 46, D497–D502 (2018).
48. Vaulou, M. et al. PeptideShaker enables reanalysis of MS-derived proteomics data sets. Nat. Biotechnol. 33, 22–24 (2015).
49. Ivanov, I. P., Firth, A. E., Michel, A. M., Atkins, J. F. & Baranov, P. V. Identification of evolutionarily conserved non-AUG-initiated N-terminal extensions in human coding sequences. Nucleic Acids Res. 39, 4220–4234 (2011).
50. Sander, J. D., Zaback, P., Joung, J. K., Voytas, D. F. & Dobbs, D. Zinc finger target (ZFIt): an engineered zinc finger/target site design tool. Nucleic Acids Res. 35, W599–W605 (2007).
51. Wilkinson, R. N., Elworthy, S., Ingham, P. W. & van Eeden, F. J. M. A method for high-throughput PCR-based genotyping of larval zebrafish tail biopsies. Biotechniques 55, 314–316 (2013).
52. Heiendil, L. et al. Creation and analysis of biochemistry constraint-based models using the CORBA Toolbox v3.0. Nat. Protoc. 14, 639–702 (2019).
53. Boutagy, N. et al. Using isolated mitochondria from minimal quantities of diseased donors and patients with dilated cardiomyopathy. J. Cell Biol. 190, 79–102 (2016).
Author contributions
L.H. and S.Z. conceptualized the study, designed, performed, and interpreted all experiments. S.Z. and G.F. carried out the SEP screen. C.L., J.H.P., C.T., R.K.C., S.J., J.C.F., and C.L.W. performed and interpreted specific experiments. B.K., V.O., G.M., and E.P. designed and performed computational analyses on sORF selection and mitochondrial functional prediction. S.N., P.S., B.Reversade, and L.S. provided reagents. R.M.S. and L.C. W. carried out proteomic analyses. C.M. carried out experiments and P.D.R. performed conservation analyses under the supervision of L.L.; N.S.J., and L.T.K. implemented MitoCore model. B.Reljic and D.A.S. performed proteomics and biochemical analyses and edited the manuscript. L.H. and S.Z. wrote the manuscript.

Competing interests
The authors declare no competing interests.

Additional information
Supplementary information is available for this paper at https://doi.org/10.1038/s41467-020-14999-2.

Correspondence and requests for materials should be addressed to L.H.

Peer review information Nature Communications thanks Cristina Ugalde and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2020