HHS Public Access

Author manuscript

J Mol Biol. Author manuscript; available in PMC 2024 November 15.

Published in final edited form as:

J Mol Biol. 2024 May 15; 436(10): 168559. doi:10.1016/j.jmb.2024.168559.

An inner mitochondrial membrane microprotein from the *SLC35A4* upstream ORF regulates cellular metabolism

Andrea L. Rocha[1], **Victor Pai**[1], **Guy Perkins**[2], **Tina Chang**[1], **Jiao Ma**[1], **Qian Chu**[1], **Joan M. Vaughan**[1], **Jolene K. Diedrich**[3], **Mark H. Ellisman**[2,*], **Alan Saghatelian**[1,*]

[1]Clayton Foundation Laboratories for Peptide Biology, Salk Institute for Biological Studies, La Jolla, CA, USA

[2]National Center for Microscopy and Imaging Research, Center for Research in Biological Systems, Department of Neurosciences, School of Medicine, University of California San Diego, La Jolla, CA, USA

[3]Mass Spectrometry Core for Proteomics and Metabolomics, The Salk Institute for Biological Studies, 10010 North Torrey Pines Road, La Jolla, CA, USA

Abstract

Upstream open reading frames (uORFs) are cis-acting elements that can dynamically regulate the translation of downstream ORFs by suppressing downstream translation under basal conditions and, in some cases, increasing translation under stress conditions. Computational and empirical methods have identified uORFs in the 5'-UTRs of approximately half of all mouse and human transcripts, making uORFs one the largest regulatory elements known. Because the prevailing dogma was that eukaryotic mRNAs produce a single functional protein, the peptides and small proteins, or microproteins, encoded by uORFs are under studied. We hypothesized that a uORF in the SLC35A4 mRNA is producing a functionalmicroprotein (SLC35A4-MP) because of its conserved amino acid sequence. Through a series of biochemical and cellular experiments, we find that the 103-amino acid SLC35A4-MP is a single-pass transmembrane inner mitochondrial membrane (IMM) microprotein. The IMM contains the protein machinery crucial for cellular respiration and ATP generation, and loss of function studies with SLC35A4-MP significantly diminish maximal cellular respiration, indicating a vital role for this microprotein in cellular metabolism. The findings add to the growing list of functional microproteins and, more generally, indicate that uORFs that encode conserved microproteins are an untapped reservoir of functional microproteins.

Keywords

upstream open reading frame (uORF); microprotein; mitochondria; inner mitochondrial membrane; cellular metabolism

[*]**Corresponding Authors:** Mark H. Ellisman and Alan Saghatelian mellisman@ucsd.edu and asaghatelian@salk.edu.

Author Contributions: A.R., T.C., M.H.E., A.S. conceived the project; A.R., T.C., G.P., and V.P. performed research and analyzed data; J.M.V. generated rabbit polyclonal SLC35A4-MP antiserum, G.P. and M.H.E. performed electron microscopy and analyzed the data, J.K.D. performed proteomics experiments, A.R., G.P., M.H.E., and A.S. wrote and edited the paper.

Competing Interest Statement: The authors declare no competing interest.

Introduction

Over half of all human mRNAs contain upstream ORFs (uORFs) [1–5]. Because the dogma in eukaryotic translation was that a single protein is translated from the longest ORF found within an mRNA, most uORFs have been overlooked for their capacity to produce proteins [6]. Instead, the biology of uORFs has focused on their role in the regulation of downstream ORF translation. Under basal conditions, uORFs inhibit downstream ORF translation by 30–80% [1, 5], but this regulation is dynamic, and under different cellular conditions uORFs can post-transcriptionally regulate the translation of the downstream ORF. The most studied example of dynamic uORF-regulated translation comes from mechanistic studies into the expression of key proteins involved in the cellular integrated stress response [7]. Specifically, uORFs are responsible for inhibiting or promoting ATF4 and CHOP during homeostasis and cell stress [8–10]. Under homeostatic conditions, translation of *ATF4* and *CHOP* uORFs represses the translation of ATF4 and CHOP by engaging the ribosome, translating the uORF sequence, and then disassociating the ribosome before it can translate the downstream ORF. However, during cell stress, phosphorylation of the translation initiation factor eIF2α leads the ribosome to bypass the uORFs to permit translation of ATF4 and CHOP proteins to counteract cell stress [8, 10–12].

In addition to their ability to regulate translation under stress conditions, uORFs have been shown to control the translation of specific mRNAs during tumorigenesis and disease. An elegant study showed that oncogenes utilize uORFs to maintain or enhance their translation while overall translation is downregulated in premalignant and cancer cells [13]. Furthermore, genetic studies have also found uORFs to have critical roles in regulating protein levels in human disease. A hereditary condition called Marie Unna hypotrichosis stems from mutations within a uORF of the human hairless homolog (HR) mRNA that leads to increased translation of HR leading to hair loss in individuals with this mutation [14]. These examples make a strong case that uORFs are a critical class of regulatory elements that control protein translation.

uORFs belong to a larger class of small open reading frames (smORFs) that encode peptides and small proteins that are collectively referred to as microproteins. A major goal of the smORF field is to determine the functions of microproteins [15, 16]. This work has led to several exciting discoveries, including the identification of novel microproteins that have explained fundamental biological mechanisms and have also revealed novel therapeutic targets [17–20]. For example, the microprotein myomixer/minion is involved in the fusion of individual myoblasts to form muscle fibers [19, 20]. Another microprotein named DWORF was shown to be a key regulator of muscle function and has been tested in preclinical gene therapy models for the treatment of cardiac hypertrophy [17, 18]. There are a handful of other examples of functionally characterized microproteins that regulate many biological processes including inflammation [21], DNA-damage repair [22], and more [23–26].

Most smORFs (≤ 100 codons) [4] that have been uncovered in the last decade via advances in proteomics (proteogenomics) and genetics (Ribo-Seq) technologies are uORFs (50–70%) [3]. However, most functional microproteins characterized to date do not come from uORFs because of the focus on uORF regulation of translation, the mistaken belief that mammalian

mRNAs produced a single (micro)protein, and the fact that until recently genetic tools that knockdown mRNAs were unable to disrupt the uORF selectively. Nevertheless, a handful of examples of functional microproteins from uORFs are known [2, 25, 27–29]. The most studied example comes from the *MIEF1* gene. The *MIEF1* uORF encodes a 70-amino acid microprotein (the MIEF-MP) that regulates the mitochondrial ribosome, mitochondrial fission, and, most recently, cell growth [2, 6, 26, 27, 29]. Characterized microproteins from uORFs represent a fraction of the total number of uORFs and additional examples are needed to better understand how to characterize uORF microproteins to understand the impact of these genetic elements in biology.

Toward this goal, we selected the microprotein derived from a uORF in the SLC35A4 5'-UTR for functional elucidation. Using biochemical fractionation, cell biology, and cellular imaging, we discovered that the SLC35A4 microprotein (SLC35A4-MP) resides in the inner mitochondrial membrane (IMM), where it regulates cellular metabolism. This work relies on the demonstration by others that CRISPR/Cas9 targeting of uORFs is a powerful method to dissect the function of individual ORFs on polycistronic mRNAs [2, 30]. Previous work has demonstrated that the SLC35A4 uORF regulates SLC35A4 translation, and our work adds to this by showing this uORF also encodes a vital mitochondrial microprotein that controls fundamental cellular biology. The bigger impact of this work comes from the realization that there are over 100 conserved uORF-encoded microproteins in a recent dataset [4] that are likely serving dual roles as regulators of downstream ORF translation as well as encoding functional microproteins like SLC35A4-MP.

Results

SLC35A4-MP is a conserved microprotein that is abundantly expressed in human cells and tissues

During the compilation of the Human Protein Atlas, a 103-amino acid microprotein translated from the *SLC35A4* mRNA uORF was first detected (SLC35A4-MP) [31]. Similar to other uORFs, such as those in *ATF4* and *CHOP* [8, 10–12], the *SLC35A4* uORF is reported to dynamically regulate the translation of the downstream ORF that encodes the SLC35A4 protein during cell stress [32]. The *SLC35A4* mRNA is spliced from three exons (Fig. 1A) with the *SLC35A4* ORF encoded entirely within exon 3 and the uORF being spliced together from exon 1, exon 2, and exon 3 (Fig. 1A). Ribosome profiling shows excellent coverage across the uORF, consistent with published data [32] (Fig. 1A), that is indicative of robust translation. While uORFs are cis-acting regulatory elements in translation, conservation analysis of the SLC35A4-MP shows that this microprotein sequence is highly conserved from humans to zebrafish (Danio rerio) (Fig. 1B), which diverged ~450 million years ago [33]. This led us to the hypothesis the SLC35A4 uORF encodes a functional microprotein, along with its role as a cis-acting regulator of translation.

Practically, the SLC35A4-MP is readily detected by proteomics in HEK293T cells indicating that it is abundant and stable in these cells to prove a tractable system to study this microprotein (Fig. 1C). Furthermore, SLC35A4 mRNA expression shows a broad expression that is indicative of a gene that regulates fundamental biology in many cell

types (Fig. 1D) and suggests that the lessons learned from studying SLC35A4 in cells can eventually be extended to animal models.

To characterize SLC35A4-MP, we began by developing rabbit polyclonal antiserum that specifically targets the N-terminal 34 amino acids of human SLC35A4-MP to overcome artifactual localization that might arise from epitope tagging and overexpression of microproteins. In parallel, we also utilized CRISPR/Cas9 to create a polyclonal SLC35A4-MP knockout (KO) HEK293T cell line for testing the antibodies (Fig. S1A). We found that the antibodies specifically label a protein at ~12 kD that is absent from cells that lack SLC35A4-MP (Fig. S1B).

Additionally, overexpression of SLC35A4-MP in HEK293T cells generated a strong signal by Western blot indicative of a robust overexpression of SLC35A4-MP (Fig. S1B) further to establish the specificity of these antibodies for SLC35A4-MP. Not surprisingly for a membrane microprotein, detection of SLC35A4-MP is amenable to the use of various cell lysis buffers containing surfactants (RIPA, RIPA + SDS, NP-40), but an acidic buffer (HCl) without surfactant gave poorer recovery yields (Fig. S1C). Collectively, these data strongly suggest that SLC35A4-MP is a stable, long-lived membrane microprotein exhibiting robust and ubiquitous expression across human cells.

SLC35A4-MP is an inner mitochondrial membrane microprotein

The structure and localization of a microprotein are necessary to help develop a functional hypothesis for these novel members of the proteome. The structure prediction tool TMHMM Server v.2.0 identified a single-pass transmembrane domain between amino acids 62 to 84 (Fig. 2A). Several of the best characterized functional microproteins contain transmembrane domains [23, 34]. Furthermore, analysis of the SLC35A4-MP sequence using BLASTP revealed that the SLC35A4-MP transmembrane domain has weak similarity to the transmembrane domain of a mitochondrial microprotein called short transmembrane mitochondrial protein 1 (STMP1) (Fig. 2B) [34]. Mitochondrial localization for SLC35A4-MP was predicted by the DeepLoc algorithm that relies on deep neural networks to confer protein localization [35] (Fig. 2C). To test the veracity of this prediction, we performed immunofluorescence experiments and biochemical fractionation of organelles to localize SLC35A4-MP (Fig. 2D and E). Confocal microscopy of endogenous SLC35A4-MP in HeLa cells (Fig. 2D) showed excellent overlap between SLC35A4-MP (shown in green in Fig. 2D) and MitoTracker (shown in red in Fig. 2D). Similarly, analysis of cytosol, nuclear, mitochondria, plasma membrane, and endoplasmic reticulum fractions from HEK293T cells by Western blot using anti-SLC35A4-MP antibodies is consistent with the imaging studies showing that the vast majority of SLC35A4-MP is in the mitochondrial fraction (Fig. 2E and S1D). This observation was also checked using polyclonal SLC35A4-MP KO and non-targeting sgRNA (hNT1) control cells, which had an SLC35A4-MP Western blot signal in the mitochondrial fraction of hNT cells but not in KO cells confirming that the detected signal is due to SLC35A4-MP and that SLC35A4-MP is localized to the mitochondria (Fig. S1D).

Mitochondria are encapsulated by two membranes, an inner mitochondrial membrane (IMM) and an outer mitochondrial membrane (OMM). To determine whether SLC35A4-MP

resides in the IMM or the OMM, we compared the effects of treating mitochondria with proteinase K while we varied the permeability of the outer membrane [24, 36, 37]. The treatment of intact mitochondria with proteinase K causes degradation of OMM proteins but leaves IMM proteins intact, whereas permeabilization of the OMM by mitochondrial swelling allows proteinase K to access and degrade IMM proteins [24, 36, 37]. As controls for this assay, we measured the effects of proteinase K treatment on the degradation of the OMM protein TOM20 and the IMM protein TIMM50 [24, 36, 37]. Immunoblotting confirmed that proteinase K treatment in intact mitochondria led to degradation of the OMM protein TOM20 but left the IMM protein TIMM50 intact, whereas proteinase K treatment while the OMM was permeabilized caused the degradation of both proteins (Fig. 2F). SLC35A4-MP showed a similar pattern to what was observed for TIMM50 (Fig. 2F), indicating that SLC35A4-MP is localized to the IMM. Immunofluorescence of MIC60, an inner mitochondrial membrane protein, and SLC35A4-MP reveal their proximity to support the biochemical localization data (Fig. S2). These data support that SLC35A4-MP is a single pass transmembrane IMM microprotein and led us to the hypothesis that SLC35A4-MP contributes to cellular energetics as this is the major function of proteins in the IMM.

SLC35A4 and SLC35A4-MP are Localized to Different Cellular Compartments

Understanding the relationship, if any, between SLC35A4 and SLC35A4-MP is critical for understanding whether these proteins are linked beyond their shared transcript. Our understanding of the biological role played by the SLC35A4 protein encoded by the main ORF is limited to a single study focused on this gene [38]. The sequence homology with other members in the SLC35 family suggests that SLC35A4 likely functions as a UDP-sugar transporter (38), which was supported by data showing that SLC35A1 and SLC35A4 are both necessary for efficient transport of cytosolic CDP-ribitol into the Golgi apparatus in HAP1 cells (38). Overexpressed and endogenous SLC35A4 localizes to the Golgi, with partial localization to the trans-Golgi, endoplasmic-reticulum–Golgi intermediate compartment (ERGIC), and endosome in the melanoma cell line (WM1341-D) [38]. Unfortunately, the antibodies used in this study are not available, and we have been unable to identify any commercial antibodies capable of detecting endogenous SLC35A4 to validate these findings in HEK293T cells (Fig. S3A). Instead, we overexpressed epitope-tagged SLC35A4-Flag and successfully detected its expression by Western blot, validating the expression and localization of the SLC35A4-Flag after transfection of HEK293T cells (Fig. 3A).

Since we were unable to use commercial antibodies to detect SLC35A4, we used mass spectrometry proteomics to try to detect the endogenous expression of this protein. We used HEK293T cells overexpressing SLC35A4-Flag to validate that we could detect SLC35A4 by proteomics, but the only peptide we detected contained a portion of the FLAG tag (Fig. S3B). Thus, we were unable to detect any peptides from the SLC35A4 sequence or validate its endogenous expression. To gather further evidence of the functional role of the SLC35A4 protein, we used TMT-proteomics to quantify proteins in cells overexpressing SLC35A4-Flag. Our findings show an increase in the number of proteins related to the endoplasmic reticulum (ER) and a specific part of the sodium:potassium-exchanging ATPase complex in these cells (Fig. 3B). To gain additional evidence for SLC35A4 Golgi-ER localization,

we immunoprecipitated SLC35A4-FLAG to identify putative interaction partners. Several ER-resident proteins were enriched upon immunoprecipitation of SLC35A4-Flag, including DHCR24, ATP6AP2, AUP1, RAB2A, SCD5, and SCD, further supporting the localization of SLC35A4 to ER (Fig. 3C). Lastly, we conducted immunofluorescence experiments in HEK293T cells expressing an either SLC35A4-FLAG or SLC35A4-Myc protein to verify that we can detect the expression of this protein and that it localizes to the Golgi-ER compartments, not the mitochondria (Fig. 3D–G). SLC35A4-FLAG and SLC35A4-Myc exhibit clear co-localization with Golgi markers, as shown by their overlap with TGN38 (Fig. 3E) and Giantin (Fig. 3F). SLC35A4-FLAG also co-localizes with the ER marker, Calnexin (Fig. 3G). By contrast, SLC35A4 does not localize with the mitochondria (Fig. 3G). In aggregate, the data indicate that SLC35A4 and SLC35A4-MP localize to different compartments and are unlikely to interact or partake in the same biology.

Generation of SLC35A4-MP-KO Clonal Cell Lines

For a more in-depth exploration into the biological function of SLC35A4-MP, we generated monoclonal CRISPR-Cas9 knockout cell lines designed to specifically target the SLC35A4 uORF. We isolated and sequenced 5 clones and found the same 1-bp insertion in clones 36 and 39, an identical 19-bp deletion in clones 29 and 35, and a 7-bp deletion in clone 35 (Fig. 4A). These indels resulted in changes to the SLC35A4-MP sequence and introduces a premature stop codon that encodes peptides that lacking a transmembrane domain (Fig. 4A). Western blot analysis of these clones showed that SLC35A4-MP expression was completely abolished in all these monoclonal cell lines (Fig. 4B). With the emergence of new stop codons, it is possible that we have impacted the stability of the SLC35A4 mRNA by triggering nonsense-mediated decay [39]. Since we cannot detect the endogenous SLC35A4, we wanted to select clones that did not change the levels of SLC35A4 mRNAs to ensure minimal perturbation of SLC35A4. qPCR analysis of the SLC35A4 mRNA (NM_080670) in each of the monoclonal KO cell lines showed insignificant changes in the mRNA in clones 29, 36, and 39 but lower levels in clones 35 and 38 (Fig. 4C). Furthermore, we imaged a hNT1 control cell line to clone 36 demonstrating the lack of SLC35A4-MP in the clonal cell line (Fig. 4D and E). Based on these results, we selected clones 29, 36, and 39 for downstream experiments.

Loss of SLC35A4-MP leads to decreased cellular respiration

The major function of proteins in the inner mitochondrial membrane is in the production of cellular energy through ATP production, which led us to examine the role of SLC35A4-MP in the regulation of cellular metabolism. To determine whether the absence in the KO cell line could have implications for cellular metabolism, we measured oxygen consumption rates (OCR) using a Seahorse XF Analyzer [40] in HEK293T clones 36, 39 and 29 that lack SLC35A4-MP (Fig 5A–E). The loss of SLC35A4-MP in these cell lines correlates with a significant decrease in basal (Fig. 5B), proton leak (Fig. 5C), maximal respiration (Fig. 5D). Furthermore, the SLC35A4-MP KO clones 39 and 29 had notably lower levels of ATP-linked OCR (Fig. 5E). These findings strongly indicate that SLC35A4-MP has a role in maintaining cellular respiratory function. One possibility is that the loss of SLC35A4-MP is leading to changes in the expression of key proteins in cellular respiration. We tested

this hypothesis by Western blot analysis of mitochondrial OXPHOS proteins but found no appreciable difference in their levels (Fig. S5A, clone 39)

Conversely, the overexpression of SLC35A4-MP (Fig. 5F–J) resulted in a modest but significant increase in maximal capacity rates (Fig. 5H), suggesting a positive impact on cellular respiration. The transfection of SLC35A4-MP into the SLC35A4-MP-KO cells (clone 36) partially rescued the phenotype, which is consistent with the notion that SLC35A4-MP is responsible for this biology (Fig. S5B–F). The reason that we do not see the complete rescue of the phenotype is that in a transient transfection not every knockout cell will receive a plasmid, so the phenotypic response is partial because we are not looking at a rescue in every cell in the experiment.

In addition, the role of SLC35A4-MP in mitochondrial membrane polarization was tested by using TMRE (tetramethyl rhodamine ethyl ester), which increases its fluorescence as a function of mitochondrial membrane polarization. While there is no appreciable difference between hNT and SLC35A4-MP-KO cells (clone 36), we observe an increase in TMRE fluorescence in cells overexpressing SLC35A4-MP (Fig. 5K–L, as well as Fig. S4). Similar results were observed with MitoTracker Green, which is also known to increase fluorescence intensity in more active and polarized mitochondria [41, 42] (Fig. S4). Based on these experiments, we conclude that SLC35A4-MP is responsible for the mitochondrial phenotypes we are observing in these cells, highlighting a heretofore unknown role for this microprotein.

SLC35A4-MP and Mitochondrial Structure

Mitochondria are a crucial component of cells, responsible for generating energy in the form of ATP. The inner mitochondrial membrane plays a significant role in this process, containing proteins that participate in the electron transport chain and ATP synthesis. Additionally, the inner membrane contributes to the formation and maintenance of cristae structures necessary for the proper functioning of mitochondria. Some inner mitochondrial membrane proteins can impact cristae structure and function, and we wanted to examine what role, if any, SLC35A4-MP has on the mitochondrial structure. We employed electron microscopy tomography that was obtained from 1.6 nm slices to evaluate the ultrastructure of cristae and cristae junction (CJ) in the mitochondria of both HEK293T hNT1 and two HEK293T SLC35A4-MP KO clonal lines (Fig. 6A–C). The mitochondria of clone 29 show a change in the mitochondrial profile area (Fig. 6D), cristae density (Fig. 6E), and CJ open size (Fig. 6F), but none of these are found in Clone 36. Based on the data, it unlikely that SLC35A4-MP has a role in mitochondrial structure, instead these subtle differences are likely due to the natural variation that would be expected from different clones from a population.

Discussion

The SLC35A4 uORF is known to regulate the translation of the SLC35A4 mRNA during cell stress [32], which is consistent with other reports that note the same phenomenon for other uORFs [43]. The consensus is that eukaryotic uORFs operate as translational regulators but are unlikely to encode functional microproteins [44]. Given that half of

all eukaryotic genes have uORFs [1], even if only a fraction of these encoded functional proteins, the numbers could be significant. We hypothesized that SLC35A4-MP is functional because its amino acid sequence is conserved (Fig. 1B) and demonstrate how understanding the cellular localization of a microprotein can develop testable hypotheses that led to its functional characterization.

Since recently identified microproteins are translated from RNAs or regions of mRNAs that were previously considered to be non-coding, such as the 5'- and 3'-UTRs, their existence changes the way we think about their encoding transcripts. For instance, the 7-kDa microprotein NoBody (NBDY) was discovered by proteogenomics as a functional microprotein product from the 'non-coding' LINC01420/LOC550643 mRNA [45]. Though there have been several studies identifying functions for LINC01420/LOC550643 mRNA in driving cancers [46, 47], the NBDY discovery indicates the mechanism is occurring via the microprotein [45] activity instead of a non-coding function for the RNA. Similarly, the discovery of a functional microprotein from the upstream smORF on SLC35A4 changes our perspective on the role of the uORF in the 5'-UTR of SLC35A4. Prior to this work, the role of this uORF was thought to act primarily as a regulator of SLC35A4 translation; however, the evidence indicates that this uORF produces a functional protein.

The relationship between uORFs and downstream ORFs has been of interest as the prevalence of uORFs in mammalian genes has grown. These mammalian uORFs indicate that polycistronic mRNAs are not limited to bacteria and overturned a dogma that mRNA from higher spies encodes for a single protein [29, 48, 49]. Operons encode several proteins that function in the same pathway, which led to the reasonable hypothesis that uORF-encoded microproteins and their downstream proteins might function in the pathway. Indeed, this model was favored when one of the first uORF microproteins characterized from uORF (AltMid51) of the mitochondrial elongation factor 1 (MIEF1 or MID51) gene was shown to have a similar function as MIEF1 [29]. Another study that defined thousands of novel smORFs by ribosome profiling, performed immunoprecipitations with ten uORF-encoded microproteins and found that half of them pulled down the protein encoded in the downstream ORF [2], putting the odds at 50% for a functional relationship between uORF microproteins and downstream proteins. This led us to test whether SLC35A4-MP and SLC35A4 potentially interact, but we found that microprotein and protein are localized to completely different subcellular compartments that are unlikely to interact. Thus, while the well characterized uORF microprotein from MIEF1 interacts with its downstream protein, the same is not true for SLC35A4-MP and SLC35A4.

In the past, we've used protein interaction partners to develop and test functional hypotheses for microproteins [22, 24, 27, 45]. However, owing to the membrane structure of this microprotein several attempts at obtaining protein interaction data with SLC35A4-MP proved to be inconsistent, with different proteins enriched in different experiments. Indeed, a recent publication that tried to functionally characterize SLC35A4-MP, mistakenly localized SLC35A4-MP to the outer mitochondrial membrane because they only used proteomics [50].

Instead, because we found that SLC35A4-MP is localized to the inner mitochondrial membrane and proteins to the inner mitochondrial membrane are linked to respiration, it allowed us to develop a testable hypothesis. To avoid any issues that could make our interpretations more challenging, we utilized CRISPR/Cas9 to create clonal cell lines that lack SLC35A4-MP and focused our biological studies on the mitochondria. Our study demonstrated the impact of SLC35A4-MP on cellular metabolism, revealing its essential role in maximal cellular respiration. By contrast, we did not observe any consistent impact of the loss of SLC35A4-MP on mitochondrial structure.

SLC35A4-MP joins a list of other microproteins including PIGBOS1 [24], SHMOOSE [51], AltMid51 [26, 29], Mitoregulin [52, 53], MOCCI [21], BRAWNIN [54], and MP31 [55] with functions in the mitochondria. Making the mitochondrial microproteins an exciting new frontier in biology. Future studies are essential to explore the *in vivo* effects of SLC35A4-MP and decipher the mechanisms that impact cellular metabolism. More generally, given that there exist hundreds of uORFs that are predicted to generate conserved microproteins [4], SLC35A4-MP is likely a forerunner of a large class of understudied by functional protein-coding genes whose characterization will enhance our understanding of microprotein biology.

Materials and Methods

SLC35A4 tissue distribution analysis

We downloaded raw RNA-Seq reads in FASTQ format from the ENCODE portal [56, 57] (https://www.encodeproject.org) for the following tissues and cell types in triplicates (see Supplementary table 1 for metadata): activated T-helper 17 cell, adrenal gland, aorta, dorsolateral prefrontal cortex, gastrocnemius medialis, heart left ventricle, heart right ventricle, kidney, lower leg skin, metanephros, mucosa of descending colon, omental fat pad, ovary, pancreas, renal cortical epithelial cell, right lobe of liver, spleen, testis, T-helper 17 cell, thyroid gland, upper lobe of left lung, and uterus. To pre-process each FASTQ file, we ran Trim Galore (v0.6.10) (https://github.com/FelixKrueger/TrimGalore) to perform quality trimming and remove adapter contamination. Afterward, we aligned the reads to the human genome (GRCh37.p13/hg19 assembly) index containing a custom GTF file from NCBI RefSeq with the reference annotation plus the appropriate exon coordinates for SLC35A4 NM_080670 using STAR (v2.5.3a) [58] with default parameters and sorted the alignments into a bam file using the software samtools (v1.13) [59]. Afterward, we performed read counting with featureCounts [60] from the package subread (v1.6.3) allowing overlapping reads (-O). We grouped the samples based on the tissues they came from. To plot the expression across the analyzed tissues, we converted the raw counts from the featureCounts output to Transcript per Million (TPM) and generated box plots to show the distribution for the replicates using the Python packages seaborn (v0.13.0) and matplotlib (v3.7.1).

Cell lines and culturing

HeLa S3 (CCL-2) and HEK293T (CRL-11268) cell lines were purchased from the American Type Culture Collection (ATCC). Both cell lines were cultured in DMEM

(Corning, 10–013-CV) supplemented with 10% fetal bovine serum (FBS) at 37 °C and 5% CO2.

Generation of an SLC35A4-MP Polyclonal Knockout Cell Line

SLC35A4-MP KO HEK293T cells were generated using two designed guide RNAs targeting exon 1 and exon 2 of the SLC35A4 genomic region (gRNAs: GGGGAAGATGGCGGATGACA and GCAGCGGCGTGTAGAAGACG, respectively). Guide RNAs were purchased from GenScript and cloned into a pSpCas9 BB-2A-Puro (PX459) vector before transfecting into HEK293T cells using Lipofectamine 2000 (Thermo Fisher Scientific, 11668–019) and Opti-MEM (Gibco, 31985–070). Following a 24 hr transfection, media was replaced with fresh DMEM supplemented with 1.25 μg/mL puromycin (SCBT, sc-108071) for selection. Media containing puromycin was replaced daily until the removal of non-transfected cells was complete. The surviving cells were validated for knockout efficiency by western blot using the rabbit polyclonal SLC35A4-MP antiserum prior to propagation.

Generation of SLC35A4-MP Monoclonal Knockout Cell Lines

Clonal SLC35A4-MP KO HEK293T cells were generated using guide RNAs targeting exon 1 of the SLC35A4 genomic region (gRNA: GGGGAAGATGGCGGATGACA). Guide RNAs were purchased from GenScript and cloned into a pSpCas9 BB-2A-Puro (PX459) vector before transfecting into HEK293T cells using Lipofectamine 2000 (Thermo Fisher Scientific, 11668–019) and Opti-MEM (Gibco, 31985–070). Following a 24 hr transfection, media was replaced with fresh DMEM supplemented with 1.25 μg/mL puromycin (SCBT, sc-108071) for selection. Media containing puromycin was replaced daily until removal of non-transfected cells was complete. Monoclonal cell lines were isolated using cell sorting by “BD Influx cell sorters” into 96 well plates. Sorted cells were spun down for 5 minutes at 500g and cultured for 1 week under normal conditions. Colonies were continually expanded from 96-well to 48-well and finally 6-well plates before collection and downstream DNA extraction with Invitrogen Purelink Genomic DNA Mini Kit according to manufacturer’s instructions. DNA was then amplified by PCR across the first exon (Forward Primer: CCACCGCCCAACTATGAACT, Reverse Primer: GGGAGCTAGGGAATGGGGTA) and was submitted for Sanger Sequencing using the forward primer (Eton Biosciences). Five clones were selected, out of which clones 36 and 39 had the same 1-bp insertion, clones 29 and 35 had an identical 19-bp deletion, and clone 35 had a 7-bp deletion. We validated cells for knockout efficiency by western blot using the rabbit polyclonal SLC35A4-MP antiserum.

Preparation of whole cell lysate for western blot

After trypsin detachment from plates cells were collected, spun down at 600g x 3 min, and washed with PBS that was preheated to 37° C. The cells were lysed by incubation with Lysis Buffer (50 mM Tris, 150 mM NaCl, 1% Triton X-100, and protease inhibitor cocktail) on ice and vortexed every 2–3 minutes. After a total of 15 minutes, the mixture was centrifuged at 13,000 g x 10 min to remove remaining cells and the supernatant was collected as the whole cell lysate. In the subsequent analysis, various cell lysis buffers, RIPA with SDS (Buffer A, 25mM Tris pH 7.6, 150 mM NaCl, 1% NP40, 1% deoxycholate,

0.1% SDS), RIPA without SDS (Buffer B, 25mM Tris pH 7.6, 150 mM NaCl, 1% NP40, 0.5% deoxycholate), 0.5% NP40 Lysis Buffer (Buffer C, 50mM Tris pH 7.6, 120 mM NaCl, 0.5% NP40, 5 mM EDTA), and HCl Lysis Buffer (Buffer D, 50 mM HCl, 0.05% Triton X-100, 0.1% 2-mercaptoethanol), were specifically employed to evaluate the efficacy of the antiserum SLC35A4-MP antibody, as shown in Supplementary Figure 1C.

Subcellular fractionation

Isolation of individual subcellular fractions from HEK293T cells was performed as described previously with slight modifications [24, 61]. HEK293T cells were homogenized by incubation with ice cold Isolation Buffer (225 mM d-mannitol, 75 mM sucrose, 0.1 mM EGTA, and 30 mM Tris-HCl pH 7.4) until most cells were lysed. The resulting homogenate was centrifuged at 600g for 5 min to collect pellet P1 and clarify the supernatant S1. Pellet P1 containing the nuclear fraction was washed three additional times with Isolation Buffer and the final pellet was resuspended in RIPA Buffer (Thermo Fisher Scientific, 89901). Supernatant S1 was then centrifuged at 7,000g for 15 min to isolate the crude mitochondrial fraction in pellet P2 from the supernatant S2. Pellet P2 was then washed with Isolation Buffer twice and the final pellet was resuspended in RIPA buffer (crude mitochondrial fraction). The supernatant S2 was centrifuged at 20,000g for 30 min to separate the plasma membrane in pellet P3 from the supernatant S3, which contains the ER and cytosolic fractions. Following two additional Isolation Buffer washes, the final pellet of P3 was resuspended in RIPA Buffer (plasma membrane fraction). S3 was further centrifuged at 100,000g for 1 h to obtain the ER (pellet P4) and cytosolic fractions. The supernatant (cytosolic fractions) was then saved while the pellet P4 was washed three additional times with Isolation Buffer. The final pellet of P4 was then resuspended in RIPA Buffer (ER fraction).

Proteinase K Protection Assay

The proteinase K protection assay was performed as described [24, 36]. HEK293T cells were homogenized by incubation with ice cold Isolation Buffer (225 mM d-mannitol, 75 mM sucrose, 0.1 mM EGTA, and 30 mM Tris-HCl pH 7.4) until most cells were lysed. The resulting homogenate was centrifuged at 600g for 5 min to collect pellet P1 and clarify the supernatant S1. Supernatant S1 was then centrifuged at 7,000g for 15 min to isolate the crude mitochondrial fraction in pellet P2 from the supernatant S2. Pellet P2 was then washed twice with an isolation buffer. Before resuspension in Isolation Buffer (225 mM mannitol, 75 mM sucrose, 50 mM HEPES pH 7.5) or in hypotonic buffer (2 mM HEPES pH 7.5) and incubated on ice for 30 min. Each buffer condition was then equally divided into two samples for a total of 4 samples; proteinase K (New England Biolabs P8107S) was added to one sample per condition at a final concentration of 50 μg/mL. Following a 30 min incubation on ice, 1 mM PMSF was added to the samples to inactivate proteinase K. The samples were then precipitated with 12% TCA and separated by SDS-PAGE followed by immunoblotting.

SDS-PAGE and immunoblotting

Samples were loaded onto 4–12% Bis-Tris gels (Thermo Fisher Scientific, NW04120BOX) and run in 1X MES-SDS buffer (Thermo Fisher Scientific, B0002) at 200 V for 27 min.

Gels were transferred onto PVDF membranes (Thermo Fisher Scientific, IB24001) using the iBlot 2 Dry Gel Transfer System (Thermo Fisher Scientific, IB21001) under the following conditions: 20 V for 1 min, 23 V for 4 min, and 25 V for 2 min. Membranes were blocked with Intercept® (TBS) Blocking Buffer (LI-COR, 927-60001) at room temperature for 1 hr prior to the addition of corresponding primary antibodies used at the indicated dilutions (Supplementary Table 2). Membranes were incubated in primary antibodies overnight on a shaker at 4°C. Upon removal of membranes from primary antibodies, membranes were washed with TBS-T (0.1% Tween-20) and incubated in Alexa-fluor labeled secondary antibodies for 1 hr (see Supplementary Table 2 for dilutions). Following TBS-T washes, images of membranes were captured using LiCor Odyssey CLx at IR700 and IR800. Subsequent analysis and quantitation were performed using ImageJ.

RT-qPCR

To isolate the total RNA, TRizol Reagent from Invitrogen was used according to the manufacturer's instructions. The SLC35A4-MP KO monoclonal cell lines and hNT1 HEK293T cells were homogenized in TRizol Reagent, mixed with chloroform in a ratio of 5:1 (v/v sample/chloroform), and then centrifuged at 12,000 x g for 15 minutes at 4°C. The total RNA was collected from the aqueous phase and precipitated by mixing with isopropanol for 15 minutes. The solution was then centrifuged at 12,000 x g for 15 minutes at 4°C. The resulting pellets were washed with 75% ethanol, dried, and resuspended in nuclease-free water. For cDNA synthesis, a High-Capacity cDNA Reverse Transcription Kit from Applied Biosystems was used. For qPCR, a 10-μL reaction mixture was prepared using Maxima SYBR Green/ROX qPCR MasterMix (2X) (ThermoFisher Scientific, K0223) and 250 nM of specific forward and reverse primers. The reactions were run in duplicate using the qPCR QuantStudio 3 qPCR (ThermoFisher Scientific) under the following conditions: 50°C −2 min, 95°C −10 min, and 40 cycles of 95°C −15s, 60°C −20s, and 72°C −30s. Dissociation protocols were conducted after every run to check for primer specificity. To obtain relative expression values, the $2^{-\Delta Ct}$ parameter was calculated for each individual sample using the Ct values of endogenous controls as a reference. The mRNA endogenous control used was 18S.

Primers:

1. mhr_18S_FW CTCAACACGGGAAACCTCAC
2. mhr_18S_RV CGCTCCACCAACTAAGAACG
3. mh_SLC35A4_FW GCTGGACCCTGATGCTACTC
4. mh_SLC35A4_RV GGGGTCCATGTAACGCTGAA

Measurements of oxygen consumption rate (OCR)

The analysis was performed using a Seahorse XFe96 Analyzer (Agilent) as described previously with slight modifications [62]. Briefly, cells were detached using trypsin (Gibco, 25200–056), resuspended in DMEM with 10% FBS, and equally divided into two collection tubes per genotype before spinning down the pellets at 600g x 3 min. For reseeding, the pellet was resuspended in DMEM with 10% FBS and then equally seeded onto a

0.01% (w/v) poly-L-lysine precoated 96-well Seahorse plate at 1×10^5 cells/well. During the reseeding, HEK293T hNT1 cells were transfected with 0.5 μg/well of pcDNA3.1+ or pSLC35A4-MP using lipofectamine 2000 for 48 hrs. After the incubation period, the media was replaced with prewarmed XF DMEM medium pH 7.4 (Agilent Technologies, 103575–100) supplemented with 10 mM D-glucose, 1 mM sodium pyruvate (Gibco, 11360–070), and 2 mM glutamax (Gibco, 35050–061). To perform a Mito stress assay, 1 μM oligomycin (Sigma, O4876), 2 μM carbonyl cyanide 4-(trifluoromethoxy)phenylhydrazone (CCCP, Sigma, 2920), and 1 μM rotenone (Sigma, R8875) and antimycin A (Sigma, A2754) were sequentially added to the assay. The assay cycle included 3 min of mixing and 3 min of measurements. Upon completion of an assay run, the cell plate was retrieved, and the cells were lysed with RIPA buffer. The protein concentrations of replicates were determined using a BCA protein assay.

TMRE experiments

HEK293T hNT1 and SLC35A4-MP KO cells were cultured on poly-L-lysine-coated coverslips (50 μg/mL, Sigma, P1399) and hNT1 cells were transfected with 2 μg/well of pSLC35A4-MP using Lipofectamine 2000 (Invitrogen, 11668019) for 48 hr. Mitochondria were stained with TMRE (tetramethylrhodamine, ethyl ester). CCCP (a mitochondrial oxidative phosphorylation uncoupler) was used as system control to a 50 μM final concentration. Cells were incubated at 37 °C for 30 min. To measure mitochondrial membrane potential, TRME solution (T669, ThermoFisher Scientific) was added to each well to a final concentration of 100 nM. Cells were washed three times with warmed PBS and then incubated in pre-warmed DMEM (Corning, 10–013-CV). Nuclei were counterstained with Hoechst 33258 (Sigma, #94403, 1:2000 in PBS), and then live images were acquired using an Olympus FV3000 Confocal Laser Scanning Microscope. For analysis of the membrane potential of mitochondria, the TMRE fluorescence intensities were measured with ImageJ software (v1.52) following published methods [50, 52].

Confocal imaging of HeLa S3 cells

HeLa S3 cells were cultured on poly-L-lysine-coated coverslips (50 μg/mL, Sigma, P1399) until reaching approximately 70% confluence. Mitochondria were stained with MitoTracker Deep Red FM (Life Technologies, M22426), followed by fixation in 4% paraformaldehyde (Polysciences, Inc., #18814) and permeabilization with 0.2% Triton X-100. Subsequently, cells were incubated overnight at 4°C with SLC35A4-MP antiserum (1:500 dilution). After washing with PBS + 0.1% Triton X-100, Alexa-488 goat anti-rabbit secondary antibody was applied for 1 hr at room temperature. Nuclei were counterstained with Hoechst 33258 (Sigma, #94403, 1:2000 in PBS), and confocal imaging was performed using a Zeiss LSM 880 Airyscan confocal microscope.

Colocalization experiments with MIC60 and MitoTracker Green by confocal imaging

HEK293T hNT1 cells were cultured on poly-L-lysine-coated coverslips (50 μg/mL, Sigma, P1399) until reaching approximately 70% confluence. MitoTracker Green FM (Invitrogen, M7514) labeled the mitochondria. Following fixation in 4% paraformaldehyde (Polysciences, Inc., #18814) and permeabilization with 0.1% Saponin, and then cells were incubated with 4% BSA for 1 hr. Subsequently, cells were incubated overnight at 4°C

with SLC35A4-MP antiserum (1:500 dilution). After washing with PBS, cells were stained with Goat TexasRed anti-rabbit secondary antibody (Invitrogen, T-2767) at 0.8 μg/mL for 1 h at room temperature (RT). MIC60 were stained with MIC60/Mitofilin Alexa 488 Conjugate (Invitrogen, MA4–56400-A488) at 5 μg/mL in PBS for 1 hr at RT. Nuclei were counterstained with Hoechst 33258 (Sigma, #94403, 1:2000 in PBS) and then imaged using Olympus FV3000 Confocal Laser Scanning Microscope. Software ImageJ was used to analysis of co-localization. The grayscale values representing green and red fluorescence along the linear distance were calculated utilizing the Plot Profile tool. Subsequently, the corresponding fluorescence channels were delineated using the rectangular tool, facilitated by Plot Profiled.

Confocal imaging of SLC35A4-Flag immunofluorescence experiments

HEK293T hNT1 cells were cultured on poly-L-lysine-coated coverslips (50 μg/mL, Sigma, P1399) and transfected with 2 μg/well of pSLC35A4 (Origene, RC200808) using Lipofectamine 2000 (Invitrogen, 11668019) for 48 hr. Mitochondria were labeled with MitoTracker Deep Red FM (Life Technologies, M22426). Following fixation in 4% paraformaldehyde (Polysciences, Inc., #18814) and permeabilization with 0.1% Saponin, cells were incubated with 4% BSA for 1 hr. Subsequently, cells were incubated overnight at 4°C with anti-FLAG (Sigma, F3165), anti-TGN38 (Santa Cruz, sc-166594), anti-Giantin (Abcam, ab37266), anti-Myc (Cell Signaling, #2276), and anti-Calnexin (ThermoScientific, PA5–34754). After washing with PBS, cells were stained with Goat TexasRed anti-rabbit and Alexa Fluor Plus 488 anti-mouse antibodies. Nuclei were counterstained with Hoechst 33258 (Sigma, #94403, 1:2000 in PBS), and imaging was performed using an Olympus FV3000 Confocal Laser Scanning Microscope. Software ImageJ was used to analysis of co-localization. The grayscale values representing green and red fluorescence along the linear distance were calculated utilizing the Plot Profile tool. Subsequently, the corresponding fluorescence channels were delineated using the rectangular tool, facilitated by Plot Profiled.

Fixation of HEK293T cells for electron microscopy (EM)

Cells were cultured in 10-cm dishes until 70% confluence. Cells were fixed with 2.5% glutaraldehyde (Electron Microscopy Sciences, 16019) and 2.5 mM calcium chloride (Thermo Fisher Scientific, BP510–500) in 0.1 M sodium cacodylate, pH 7.4 (Electron Microscopy Sciences, 12300) for 5 min followed by another 60 min on ice. Following 0.1 M sodium cacodylate buffer pH 7.4 rinses, cells were blocked in 10 mM glycine (VWR International, 470301–176) and 10 mM potassium cyanide (Sigma-Aldrich, 60178) for 15 min on ice and washed an additional 3 times with 0.1 M sodium cacodylate buffer. A solution of 3% potassium ferrocyanide (Acros Organics, 424130050) in 0.3 M cacodylate buffer and 4 mM calcium chloride was combined with an equal volume of 4% (wt/wt) aqueous osmium tetroxide (Electron Microscopy Sciences, 19150) then added to the samples and incubated for 30 min on ice. After rinsing the cells in cold double deionized water (DDW), the cells were placed in 2% aqueous uranyl acetate for 1 hr. The cells were then rinsed in cold DDW before dehydration series using 20%, 30%, 50%, 70%, 90%, 100% ethanol. A Durcupan ACM resin was prepared as follows: 11 g Part A (Sigma-Aldrich, 44611), 10 g Part B (Sigma-Aldrich, 44612), 0.3 g Part C (Sigma-Aldrich, 44613), and 0.1 g Part D (Sigma-Aldrich, 44614). The samples were placed into 50% solution of Durcupan

in ethanol for 30 min at room temperature followed by 100% Durcupan for 30 min. The resulting sample was then placed at 60°C for 48 hr before analysis.

Transmission electron microscopy

Thin sections about 70 nm thick were cut from the cell blocks with a Leica ultramicrotome and placed on 200-mesh uncoated thin-bar copper grids. A Tecnai Spirit (FEI; Hillsboro, OR) electron microscope operated at 80 kV was used to record images with a Gatan 2Kx2K CCD camera at 2.88 nm/pixel. ImageJ was used to measure the number of mitochondria and cytoplasmic areas.

EM tomography

Semi-thick sections of thickness about 400 nm were cut from the cell blocks with a Leica ultramicrotome and placed on 200-mesh uncoated thin-bar copper grids. 20-nm colloidal gold particles were deposited on each side of the grid to serve as fiducial cues. The grids were irradiated for about 15 min before initiating a double-tilt series to limit anisotropic specimen thinning during image collection. During data collection, the illumination was held to near parallel beam conditions and the beam intensity was held constant. Tilt series were captured using SerialEM (University of Colorado, Boulder, CO) software on a Tecnai High Base Titan (FEI; Hillsboro, OR) electron microscope operated at 300 kV and 0.81 nm/pixel. Images were recorded with a Gatan 4Kx4K CCD camera. Each double-tilt series consisted of first collecting 121 images taken at 1-degree increment over a range of −60 to +60 degrees followed by rotating the grid 90 degrees and collecting another 121 images with the same tilt increment. To improve the signal-to-noise ratio, 2x binning was performed on each image by averaging a 2×2 x-y pixel box into 1 pixel.

The IMOD package (https://en.wikipedia.org/wiki/IMOD_(software)) was used for alignment, reconstruction, and volume segmentation. R-weighted back projection was used to generate the reconstructions. Volume segmentation of the mitochondrial outer membrane was performed using IMOD by tracing in each of the 1.62 nm-thick x-y planes that the object appeared, which then created stacks of contours with the Drawing Tools plug-in in IMOD. The traced contours were then surface rendered by turning contours into meshes to generate a 3D model. The surface-rendered volumes were visualized using 3DMOD. Measurements of crista junction width, height, and number, crista number, and mitochondrial outer membrane surface area were made within tomographic volumes following volume segmentation and reported using IMODinfo. Crista density, defined as the total membrane surface area of the cristae divided by the volume of the mitochondrion, was measured using the ImageJ stereology plug-in on slices from tomographic volumes.

Antisera production-animal care

All animal procedures were approved by the Institutional Animal Care and Use Committee of the Salk Institute and were conducted in accordance with the PHS Policy on Humane Care and Use of Laboratory Animals (PHS Policy, 2015), the U.S. Government Principles for Utilization and Care of Vertebrate Animals Used in Testing, Research and Training, the NRC Guide for Care and Use of Laboratory Animals (8th edition) and the USDA Animal Welfare Act and Regulations. All animals were housed in an AAALAC accredited facility

in a climate-controlled environment (65–72 degrees Fahrenheit, 30–70% humidity) under 12 light/12 dark cycles. Upon arrival, animals were physically examined by veterinary staff for good health and acclimated for at least two weeks prior to initiation of antisera production. Each animal was monitored daily by the veterinary staff for signs of complications and weighed every two weeks. Routine physical exams were also performed by the veterinarian quarterly on all rabbits. For production of antisera against human SLC35A4-MP, three 10- to 12-week-old, female New Zealand white rabbits, weighing 3.0 to 3.2 kg at beginning of the study, were procured from Irish Farms (I.F.P.S. Inc., Norco, California, USA). Rabbits were provided with ad libitum feed (5326 Lab Diet High Fiber), micro-filtered water and weekly fruits and vegetables and alfalfa hay for enrichment.

Generation of SLC35A4-MP antisera

Antisera against SLC35A4-MP was raised in rabbits against a synthetic peptide fragment encoding human Cys34 SLC35A4(2–34) coupled to maleimide activated keyhole limpet hemocyanin (Thermo Fisher Scientific) per manufacturer's instructions. The SLC35A4-MP peptide, ADDKDSLPKLKDLAFLKNQLESLQRRVEDEVNC, was synthesized, C18 HPLC purified to 99%, and amino acid sequence verified using mass spectrometry by RS Synthesis (Louisville, KY). The immunogen was prepared by emulsification of Freund's complete adjuvant-modified Mycobacterium butyricum (EMD Millipore) with an equal volume of phosphate buffered saline (PBS) containing 1.0 mg conjugate/ml for the first two injections. For booster injections, incomplete Freund's adjuvant was mixed with an equal volume of PBS containing 0.5 mg conjugate/ml. For each immunization, an animal received a total of 1 ml of emulsion in 20 intradermal sites in the lumbar region, 0.5 mg total protein conjugate for the first two injections and 0.25 mg total protein conjugate for all subsequent booster injections. Three individual rabbits were injected every three weeks and were bled one week following booster injections, <10% total blood volume. Rabbits were administered 1–2 mg/kg Acepromazine IM prior to injections of antigen or blood withdrawal. At the termination of study, rabbits were exsanguinated under anesthesia (ketamine 50 mg/kg and aceprozamine 1 mg/kg, IM) and euthanized with an overdose of pentobarbital sodium and phenytoin sodium (1 ml/4.5 kg of body weight IC to effect). After blood was collected the death of animals was confirmed. All animal procedures were conducted by experienced veterinary technicians, under the supervision of Salk Institute veterinarians.

Characterization and purification of SLC35A4-MP antisera

Each bleed from each animal was tested at multiple doses for the ability to recognize the synthetic peptide antigen; bleeds with highest titers were further analyzed by western immunoblot for the ability to recognize the full-length endogenous SLC35A4-MP and to check for cross-reactivity to other proteins. Antisera with the best characteristics of titer against the synthetic peptide antigen, ability to recognize the endogenous protein, and specificity were antigen affinity purified and used for all studies. Rabbit PBL#7383 was purified using human Cys34 SLC35A4-MP(2–34) coupled to SulfoLink coupling resin (Thermo Fisher Scientific Cat #20401), 2 mg peptide on 2 ml bed volume resin, per manufacturer's instructions. To ensure that the same batch of purified antibody could be used for this study, a large volume, ~ 20 ml sera pooled from bleeds with similar profiles, was purified. The pooled serum was mixed with an equal volume of PBS, filtered through

a 5 μm syringe filter, and then rotated overnight with SLC35A4-MP(2–34)-SulfoLink resin. Unbound material was collected using a column, the resin was washed four times, using 2 column volumes each time, with 50 mM NaHepes, 100 mM NaCl, 0.1% Triton X-100, pH 7.5 followed by washing three times, using 2 column volumes each time, with 100 mM NaCitrate pH 4.5. Finally, the column was eluted with 1 N acetic acid and fractions of 1 column volume each were collected. The fractions were monitored for protein and the immunoglobulins containing fractions were pooled and dialyzed against phosphate buffered saline at 4°C using a Slide-A-Lyzer cassette MWCO 7K (Thermo Fisher Scientific).

Preparation of MS samples and instrumentation

Trypsin digestion and sample preparation for MS/MS was performed as previously (7). Briefly, samples were precipitated with 6.1 N trichloroacetic acid (TCA) (MP Biomedicals, 196057) to reach a final TCA concentration of 23% (v/v) and incubated on ice overnight. Protein precipitate was washed with ice cold acetone twice before dissolving in 8 M urea (Alfa Aesar, 36428) followed by reduction with tris(2-carboxyethyl)phosphine hydrochloride (TCEP, Thermo Fisher Scientific, 20491, 5 mM final concentration) then alkylation with 2-chloroacetamide (TCI, C2536, 10 mM final concentration). Upon addition of triethylammonium bicarbonate buffer (TEAB, Sigma, T7408, final urea concentration is 2 M), samples were digested overnight at 37 °C in the dark by trypsin (Promega, V5111). The reaction was quenched with formic acid to achieve a final concentration of 5% (v/v). The digested sample was then centrifuged at 15,000g for 15 min and the supernatant was collected for further analysis on a Q Exactive mass spectrometer (Thermo Fisher Scientific).

TMT-labeling proteomics protocol

Samples were precipitated by methanol/ chloroform and redissolved in 8 M urea/100 mM TEAB, pH 8.5. Proteins were reduced with 5 mM tris(2-carboxyethyl)phosphine hydrochloride (TCEP, Sigma-Aldrich) and alkylated with 10 mM chloroacetamide (Sigma-Aldrich). Proteins were digested overnight at 37 °C in 2 M urea/100 mM TEAB, pH 8.5, with trypsin (Promega). The digested peptides were labeled with 10-plex TMT (Thermo product 90309, lot XB318561), pooled samples were fractionated by basic reversed phase (Thermo 84868).

The TMT labeled samples were analyzed on a Orbitrap Eclipse Tribrid mass spectrometer (Thermo). Samples were injected directly onto a 25 cm, 100 μm ID column packed with BEH 1.7 μm C18 resin (Waters). Samples were separated at a flow rate of 300 nL/min on an EasynLC 1200 (Thermo). Buffer A and B were 0.1% formic acid in water and 90% acetonitrile, respectively. A gradient of 1–10% B over 30 min, an increase to 35% B over 120 min, an increase to 100% B over 20 min and held at 100% B for a 10 min was used for a 180 min total run time.

Peptides were eluted directly from the tip of the column and nanosprayed directly into the mass spectrometer by application of 2.5 kV voltage at the back of the column. The Eclipse was operated in a data dependent mode. Full MS1 scans were collected in the Orbitrap at 120k resolution. The cycle time was set to 3 s, and within this 3 s the most abundant ions per scan were selected for CID MS/MS in the ion trap. MS3 analysis with multinotch isolation

(SPS3) was utilized for detection of TMT reporter ions at 60k resolution [63]. Monoisotopic precursor selection was enabled and dynamic exclusion was used with exclusion duration of 60 s.

Protein and peptide identification were done with Integrated Proteomics Pipeline – IP2 (Integrated Proteomics Applications). Tandem mass spectra were extracted from raw files using RawConverter [64] and searched with ProLuCID [65] against Uniprot human database. The search space included all fully-tryptic and half-tryptic peptide candidates. Carbamidomethylation on cysteine and TMT on lysine and peptide N-term were considered as static modifications. Data was searched with 50 ppm precursor ion tolerance and 600 ppm fragment ion tolerance. Identified proteins were filtered to using DTASelect [66] and utilizing a target-decoy database search strategy to control the false discovery rate to 1% at the protein level [67]. Quantitative analysis of TMT was done with Census [68] filtering reporter ions with 10 ppm mass tolerance and 0.6 isobaric purity filter.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank members of the Saghatelian and Ellisman labs for the invaluable comments and suggestions received throughout our study. This research was supported by Frederik Paulsen Chair (A.S.), NIH (P30 CA014195, R01GM102491, A.S.), as well as grants to MHE (NIH/NIAR01 AG081037, NIH/NIMHRF 1MH129261, NIH U24NS120055, NIH R01NS108934, NIH/NIAR01 AG062479, NIHR01GM138780, NSF 2014862, NIH R01AG065549, NIH S10OD021784).

References

[1]. Calvo SE, Pagliarini DJ, Mootha VK. Upstream open reading frames cause widespread reduction of protein expression and are polymorphic among humans. Proc Natl Acad Sci U S A. 2009;106:7507–12. [PubMed: 19372376]

[2]. Chen J, Brunner AD, Cogan JZ, Nunez JK, Fields AP, Adamson B, et al. Pervasive functional translation of noncanonical human open reading frames. Science. 2020;367:1140–6. [PubMed: 32139545]

[3]. Ingolia NT, Lareau LF, Weissman JS. Ribosome profiling of mouse embryonic stem cells reveals the complexity and dynamics of mammalian proteomes. Cell. 2011;147:789–802. [PubMed: 22056041]

[4]. Martinez TF, Chu Q, Donaldson C, Tan D, Shokhirev MN, Saghatelian A. Accurate annotation of human protein-coding small open reading frames. Nat Chem Biol. 2020;16:458–68. [PubMed: 31819274]

[5]. Barbosa C, Peixeiro I, Romao L. Gene expression regulation by upstream open reading frames and human disease. PLoS Genet. 2013;9:e1003529. [PubMed: 23950723]

[6]. Brunet MA, Levesque SA, Hunting DJ, Cohen AA, Roucou X. Recognition of the polycistronic nature of human genes is critical to understanding the genotype-phenotype relationship. Genome Res. 2018;28:609–24. [PubMed: 29626081]

[7]. Boye E, Grallert B. eIF2alpha phosphorylation and the regulation of translation. Curr Genet. 2020;66:293–7. [PubMed: 31485739]

[8]. Harding HP, Novoa I, Zhang Y, Zeng H, Wek R, Schapira M, Ron D. Regulated translation initiation controls stress-induced gene expression in mammalian cells. Mol Cell. 2000;6:1099–108. [PubMed: 11106749]

[9]. Jousse C, Bruhat A, Carraro V, Urano F, Ferrara M, Ron D, Fafournoux P. Inhibition of CHOP translation by a peptide encoded by an open reading frame localized in the chop 5'UTR. Nucleic Acids Res. 2001;29:4341–51. [PubMed: 11691921]

[10]. Young SK, Wek RC. Upstream Open Reading Frames Differentially Regulate Gene-specific Translation in the Integrated Stress Response. J Biol Chem. 2016;291:16927–35. [PubMed: 27358398]

[11]. Palam LR, Baird TD, Wek RC. Phosphorylation of eIF2 facilitates ribosomal bypass of an inhibitory upstream ORF to enhance CHOP translation. J Biol Chem. 2011;286:10939–49. [PubMed: 21285359]

[12]. Vattem KM, Wek RC. Reinitiation involving upstream ORFs regulates ATF4 mRNA translation in mammalian cells. Proc Natl Acad Sci U S A. 2004;101:11269–74. [PubMed: 15277680]

[13]. Sendoel A, Dunn JG, Rodriguez EH, Naik S, Gomez NC, Hurwitz B, et al. Translation from unconventional 5' start sites drives tumour initiation. Nature. 2017;541:494–9. [PubMed: 28077873]

[14]. Wen Y, Liu Y, Xu Y, Zhao Y, Hua R, Wang K, et al. Loss-of-function mutations of an inhibitory upstream ORF in the human hairless transcript cause Marie Unna hereditary hypotrichosis. Nat Genet. 2009;41:228–33. [PubMed: 19122663]

[15]. Ji X, Cui C, Cui Q. smORFunction: a tool for predicting functions of small open reading frames and microproteins. BMC Bioinformatics. 2020;21:455. [PubMed: 33054771]

[16]. Saghatelian A, Couso JP. Discovery and characterization of smORF-encoded bioactive polypeptides. Nat Chem Biol. 2015;11:909–16. [PubMed: 26575237]

[17]. Makarewich CA, Bezprozvannaya S, Gibson AM, Bassel-Duby R, Olson EN. Gene Therapy With the DWORF Micropeptide Attenuates Cardiomyopathy in Mice. Circ Res. 2020;127:1340–2. [PubMed: 32878549]

[18]. Makarewich CA, Munir AZ, Schiattarella GG, Bezprozvannaya S, Raguimova ON, Cho EE, et al. The DWORF micropeptide enhances contractility and prevents heart failure in a mouse model of dilated cardiomyopathy. Elife. 2018;7.

[19]. Ramirez-Martinez A, Zhang Y, van den Boogaard MJ, McAnally JR, Rodriguez-Caycedo C, Chai AC, et al. Impaired activity of the fusogenic micropeptide Myomixer causes myopathy resembling Carey-Fineman-Ziter syndrome. J Clin Invest. 2022;132.

[20]. Zhang Q, Vashisht AA, O'Rourke J, Corbel SY, Moran R, Romero A, et al. The microprotein Minion controls cell fusion and muscle formation. Nat Commun. 2017;8:15664. [PubMed: 28569745]

[21]. Lee CQE, Kerouanton B, Chothani S, Zhang S, Chen Y, Mantri CK, et al. Coding and non-coding roles of MOCCI (C15ORF48) coordinate to regulate host inflammation and immunity. Nat Commun. 2021;12:2130. [PubMed: 33837217]

[22]. Arnoult N, Correia A, Ma J, Merlo A, Garcia-Gomez S, Maric M, et al. Regulation of DNA repair pathway choice in S and G2 phases by the NHEJ inhibitor CYREN. Nature. 2017;549:548–52. [PubMed: 28959974]

[23]. Anderson DM, Makarewich CA, Anderson KM, Shelton JM, Bezprozvannaya S, Bassel-Duby R, Olson EN. Widespread control of calcium signaling by a family of SERCA-inhibiting micropeptides. Sci Signal. 2016;9:ra119. [PubMed: 27923914]

[24]. Chu Q, Martinez TF, Novak SW, Donaldson CJ, Tan D, Vaughan JM, et al. Regulation of the ER stress response by a mitochondrial microprotein. Nat Commun. 2019;10:4883. [PubMed: 31653868]

[25]. Cloutier P, Poitras C, Faubert D, Bouchard A, Blanchette M, Gauthier M-S, Coulombe B. Upstream ORF-Encoded ASDURF Is a Novel Prefoldin-like Subunit of the PAQosome. Journal of Proteome Research. 2020;19:18–27. [PubMed: 31738558]

[26]. Delcourt V, Brunelle M, Roy AV, Jacques JF, Salzet M, Fournier I, Roucou X. The Protein Coded by a Short Open Reading Frame, Not by the Annotated Coding Sequence, Is the Main Gene Product of the Dual-Coding Gene MIEF1. Mol Cell Proteomics. 2018;17:2402–11. [PubMed: 30181344]

[27]. Rathore A, Chu Q, Tan D, Martinez TF, Donaldson CJ, Diedrich JK, et al. MIEF1 Microprotein Regulates Mitochondrial Translation. Biochemistry. 2018;57:5564–75. [PubMed: 30215512]

[28]. Rensvold JW, Shishkova E, Sverchkov Y, Miller IJ, Cetinkaya A, Pyle A, et al. Defining mitochondrial protein functions through deep multiomic profiling. Nature. 2022.

[29]. Samandi S, Roy AV, Delcourt V, Lucier JF, Gagnon J, Beaudoin MC, et al. Deep transcriptome annotation enables the discovery and functional characterization of cryptic small proteins. Elife. 2017;6.

[30]. Prensner JR, Enache OM, Luria V, Krug K, Clauser KR, Dempster JM, et al. Noncanonical open reading frames encode functional proteins essential for cancer cell survival. Nat Biotechnol. 2021;39:697–704. [PubMed: 33510483]

[31]. Kim MS, Pinto SM, Getnet D, Nirujogi RS, Manda SS, Chaerkady R, et al. A draft map of the human proteome. Nature. 2014;509:575–81. [PubMed: 24870542]

[32]. Andreev DE, O'Connor PB, Fahey C, Kenny EM, Terenin IM, Dmitriev SE, et al. Translation of 5' leaders is pervasive in genes resistant to eIF2 repression. Elife. 2015;4:e03971. [PubMed: 25621764]

[33]. Liu TX, Zhou Y, Kanki JP, Deng M, Rhodes J, Yang HW, et al. Evolutionary conservation of zebrafish linkage group 14 with frequently deleted regions of human chromosome 5 in myeloid malignancies. Proc Natl Acad Sci U S A. 2002;99:6136–41. [PubMed: 11983906]

[34]. Zhang D, Xi Y, Coccimiglio ML, Mennigen JA, Jonz MG, Ekker M, Trudeau VL. Functional prediction and physiological characterization of a novel short trans-membrane protein 1 as a subunit of mitochondrial respiratory complexes. Physiol Genomics. 2012;44:1133–40. [PubMed: 23073385]

[35]. Almagro Armenteros JJ, Sonderby CK, Sonderby SK, Nielsen H, Winther O. DeepLoc: prediction of protein subcellular localization using deep learning. Bioinformatics. 2017;33:3387–95. [PubMed: 29036616]

[36]. Aerts L, Craessaerts K, De Strooper B, Morais VA. PINK1 kinase catalytic activity is regulated by phosphorylation on serines 228 and 402. J Biol Chem. 2015;290:2798–811. [PubMed: 25527497]

[37]. Dimmer KS, Navoni F, Casarin A, Trevisson E, Endele S, Winterpacht A, et al. LETM1, deleted in Wolf-Hirschhorn syndrome is required for normal mitochondrial morphology and cellular viability. Hum Mol Genet. 2008;17:201–14. [PubMed: 17925330]

[38]. Sosicka P, Maszczak-Seneczko D, Bazan B, Shauchuk Y, Kaczmarek B, Olczak M. An insight into the orphan nucleotide sugar transporter SLC35A4. Biochim Biophys Acta Mol Cell Res. 2017;1864:825–38. [PubMed: 28167211]

[39]. Perrin-Vidoz L, Sinilnikova OM, Stoppa-Lyonnet D, Lenoir GM, Mazoyer S. The nonsense-mediated mRNA decay pathway triggers degradation of most BRCA1 mRNAs bearing premature termination codons. Hum Mol Genet. 2002;11:2805–14. [PubMed: 12393792]

[40]. Divakaruni AS, Rogers GW, Murphy AN. Measuring Mitochondrial Function in Permeabilized Cells Using the Seahorse XF Analyzer or a Clark-Type Oxygen Electrode. Curr Protoc Toxicol. 2014;60:25 2 1–16.

[41]. Doherty E, Perl A. Measurement of Mitochondrial Mass by Flow Cytometry during Oxidative Stress. React Oxyg Species (Apex). 2017;4:275–83. [PubMed: 29806036]

[42]. Little AC, Kovalenko I, Goo LE, Hong HS, Kerk SA, Yates JA, et al. High-content fluorescence imaging with the metabolic flux assay reveals insights into mitochondrial properties and functions. Commun Biol. 2020;3:271. [PubMed: 32472013]

[43]. Moro SG, Hermans C, Ruiz-Orera J, Alba MM. Impact of uORFs in mediating regulation of translation in stress conditions. BMC Mol Cell Biol. 2021;22:29. [PubMed: 33992089]

[44]. Zhang H, Wang Y, Wu X, Tang X, Wu C, Lu J. Determinants of genome-wide distribution and evolution of uORFs in eukaryotes. Nat Commun. 2021;12:1076. [PubMed: 33597535]

[45]. D'Lima NG, Ma J, Winkler L, Chu Q, Loh KH, Corpuz EO, et al. A human microprotein that interacts with the mRNA decapping complex. Nat Chem Biol. 2017;13:174–80. [PubMed: 27918561]

[46]. Wu HF, Lu TJ, Lo YH, Tu YT, Chen YR, Lee MC, et al. Long Noncoding RNA LOC550643 Acts as an Oncogene in the Growth Regulation of Colorectal Cancer Cells. Cells. 2022;11.

[47]. Tsai KW, Chong KH, Li CH, Tu YT, Chen YR, Lee MC, et al. LOC550643, a Long Non-coding RNA, Acts as Novel Oncogene in Regulating Breast Cancer Growth and Metastasis. Front Cell Dev Biol 2021;9:695632. [PubMed: 34354991]

[48]. Mouilleron H, Delcourt V, Roucou X. Death of a dogma: eukaryotic mRNAs can code for more than one protein. Nucleic Acids Res. 2016;44:14–23. [PubMed: 26578573]

[49]. Blumenthal T Operons in eukaryotes. Brief Funct Genomic Proteomic. 2004;3:199–211. [PubMed: 15642184]

[50]. Yang H, Li Q, Stroup EK, Wang S, Ji Z. Widespread stable noncanonical peptides identified by integrated analyses of ribosome profiling and ORF features. Nat Commun. 2024;15:1932. [PubMed: 38431639]

[51]. Miller B, Kim SJ, Mehta HH, Cao K, Kumagai H, Thumaty N, et al. Mitochondrial DNA variation in Alzheimer's disease reveals a unique microprotein called SHMOOSE. Mol Psychiatry. 2023;28:1813–26. [PubMed: 36127429]

[52]. Stein CS, Jadiya P, Zhang X, McLendon JM, Abouassaly GM, Witmer NH, et al. Mitoregulin: A lncRNA-Encoded Microprotein that Supports Mitochondrial Supercomplexes and Respiratory Efficiency. Cell Rep. 2018;23:3710–20 e8. [PubMed: 29949756]

[53]. Zhang S, Guo Y, Fidelito G, Robinson DRL, Liang C, Lim R, et al. LINC00116-encoded microprotein mitoregulin regulates fatty acid metabolism at the mitochondrial outer membrane. iScience. 2023;26:107558. [PubMed: 37664623]

[54]. Zhang S, Reljic B, Liang C, Kerouanton B, Francisco JC, Peh JH, et al. Mitochondrial peptide BRAWNIN is essential for vertebrate respiratory complex III assembly. Nat Commun. 2020;11:1312. [PubMed: 32161263]

[55]. Huang N, Li F, Zhang M, Zhou H, Chen Z, Ma X, et al. An Upstream Open Reading Frame in Phosphatase and Tensin Homolog Encodes a Circuit Breaker of Lactate Metabolism. Cell Metab. 2021;33:128–44 e9. [PubMed: 33406399]

[56]. Luo Y, Hitz BC, Gabdank I, Hilton JA, Kagda MS, Lam B, et al. New developments on the Encyclopedia of DNA Elements (ENCODE) data portal. Nucleic Acids Res. 2020;48:D882–D9. [PubMed: 31713622]

[57]. Consortium EP. An integrated encyclopedia of DNA elements in the human genome. Nature. 2012;489:57–74. [PubMed: 22955616]

[58]. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: ultrafast universal RNA-seq aligner. Bioinformatics. 2013;29:15–21. [PubMed: 23104886]

[59]. Danecek P, Bonfield JK, Liddle J, Marshall J, Ohan V, Pollard MO, et al. Twelve years of SAMtools and BCFtools. Gigascience. 2021;10.

[60]. Liao Y, Smyth GK, Shi W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics. 2014;30:923–30. [PubMed: 24227677]

[61]. Wieckowski MR, Giorgi C, Lebiedzinska M, Duszynski J, Pinton P. Isolation of mitochondria-associated membranes and mitochondria from animal tissues and cells. Nat Protoc. 2009;4:1582–90. [PubMed: 19816421]

[62]. Gu X, Ma Y, Liu Y, Wan Q. Measurement of mitochondrial respiration in adherent cells by Seahorse XF96 Cell Mito Stress Test. STAR Protoc. 2021;2:100245. [PubMed: 33458707]

[63]. McAlister GC, Nusinow DP, Jedrychowski MP, Wuhr M, Huttlin EL, Erickson BK, et al. MultiNotch MS3 enables accurate, sensitive, and multiplexed detection of differential expression across cancer cell line proteomes. Anal Chem. 2014;86:7150–8. [PubMed: 24927332]

[64]. He L, Diedrich J, Chu YY, Yates JR 3rd. Extracting Accurate Precursor Information for Tandem Mass Spectra by RawConverter. Anal Chem. 2015;87:11361–7. [PubMed: 26499134]

[65]. Xu T, Park SK, Venable JD, Wohlschlegel JA, Diedrich JK, Cociorva D, et al. ProLuCID: An improved SEQUEST-like algorithm with enhanced sensitivity and specificity. J Proteomics. 2015;129:16–24. [PubMed: 26171723]

[66]. Tabb DL, McDonald WH, Yates JR 3rd. DTASelect and Contrast: tools for assembling and comparing protein identifications from shotgun proteomics. J Proteome Res. 2002;1:21–6. [PubMed: 12643522]

[67]. Peng J, Elias JE, Thoreen CC, Licklider LJ, Gygi SP. Evaluation of multidimensional chromatography coupled with tandem mass spectrometry (LC/LC-MS/MS) for large-scale protein analysis: the yeast proteome. J Proteome Res. 2003;2:43–50. [PubMed: 12643542]

[68]. Park SK, Aslanian A, McClatchy DB, Han X, Shah H, Singh M, et al. Census 2: isobaric labeling data analysis. Bioinformatics. 2014;30:2208–9. [PubMed: 24681903]

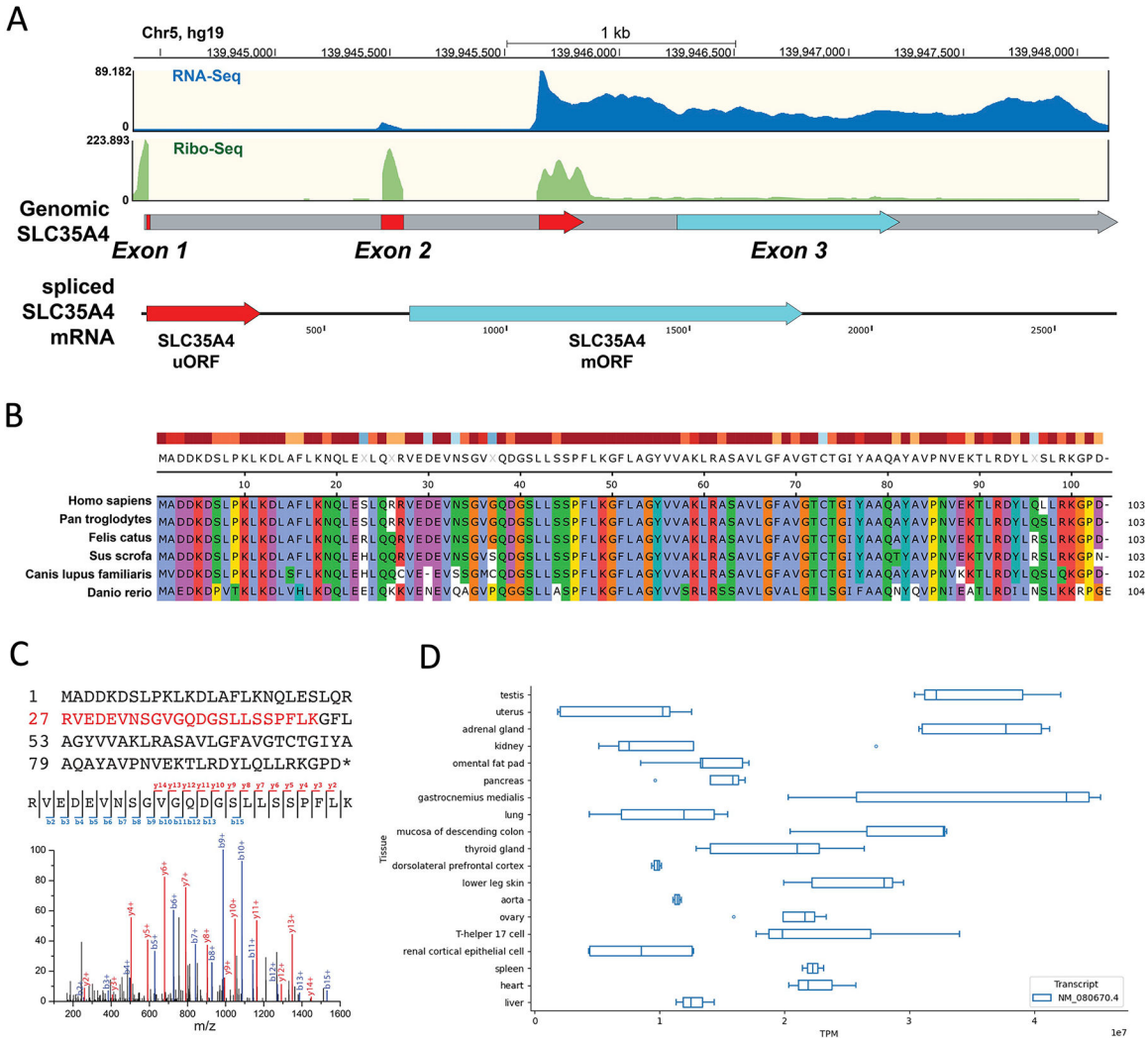

Fig. 1. Exploring the spatial interactions and functional significance of microprotein SLC35A4-MP.

(A) The human *SLC35A4* genomic locus on chromosome 5 shows mRNA expression (RNA-Seq, blue) and ribosome coverage on the mRNA (Ribo-Seq, green). The *SLC35A4* uORF (red) is distributed across all three *SLC35A4* exons and is fully formed after splicing, whereas the *SLC35A4* mORF (turquoise) is contained entirely within Exon 3. This figure captured the spatial dynamics of genetic elements in SLC35A4. (B) SLC35A4-MP is highly conserved, suggesting that the microprotein, not just the uORF, is functional. The colors indicate the properties of individual amino acid residues (blue, hydrophobic; purple, negative charge; red, positive charge; and green, polar). (C) The *SLC35A4* mRNA contains a uORF that produces a microprotein, SLC35A4-MP, that is detected by proteomics (an MS/MS spectrum of a unique tryptic peptide fragment). (D) A box plot displays the expression of SLC35A4 mRNA in various human tissues and cells. The plot shows different quartiles for the TPMs (transcripts per million). For this analysis, we used the exon coordinates for transcripts that contain SLC35A4 (NM_080670.4).

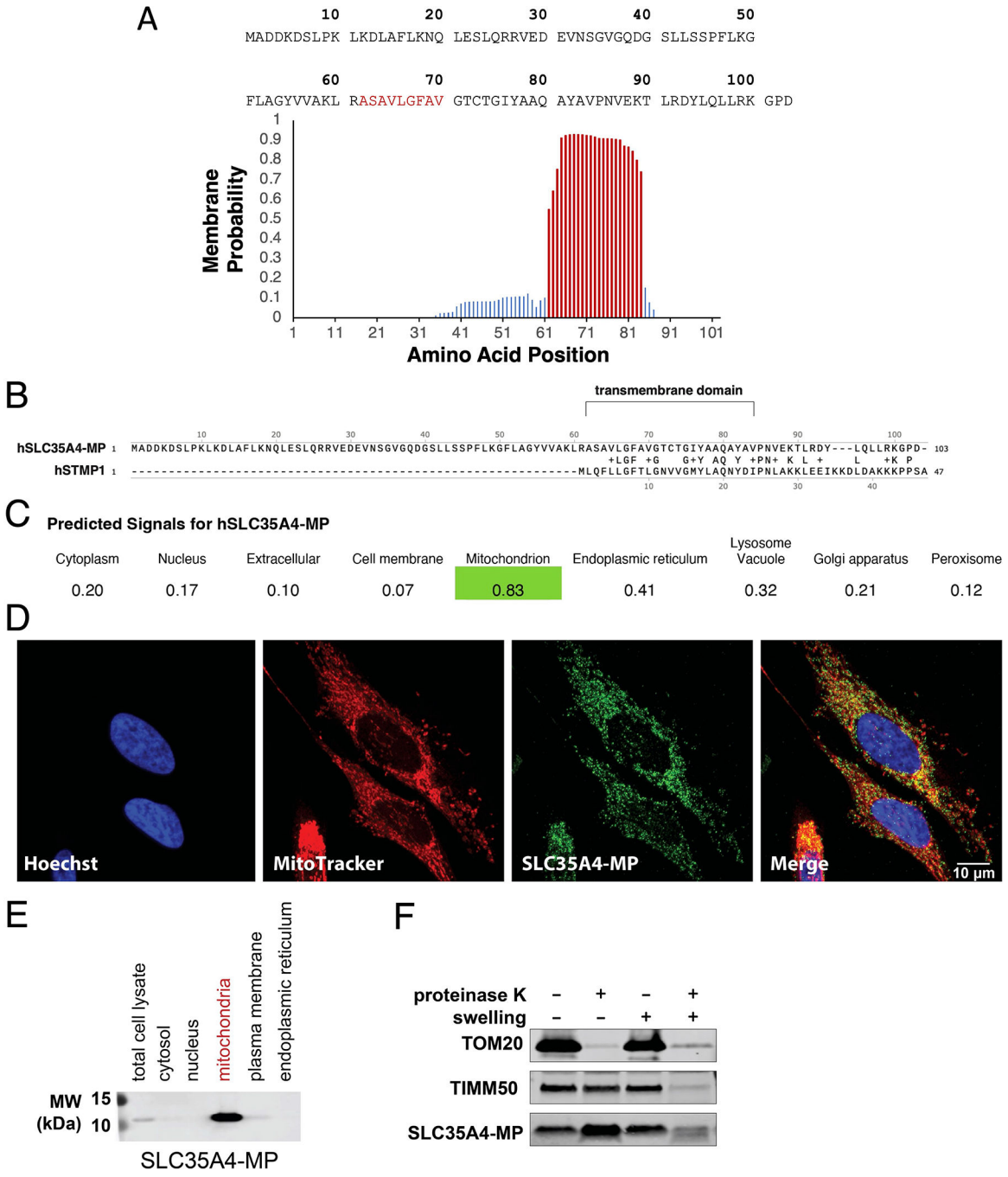

Fig. 2. SLC35A4-MP is a single-pass transmembrane protein that is localized to the inner mitochondrial membrane.

(A) Analysis of the SLC35A4-MP amino acid sequence using TMHMM Server identifies a single-pass transmembrane domain (TMD) consisting of amino acid residues 62 through 84 highlighted in red. (B) Sequence alignment of human SLC35A4-MP and STMP1. The alignment of human SLC35A4-MP (103 amino acids) and human STMP1 (47 amino acids) illustrates the two proteins share sequence similarity, particularly in the SLC35A4-MP transmembrane domain. (C) Algorithmic prediction of mitochondrial localization of SLC35A4-MP using DeepLoc. (D) Immunofluorescence of fixed HeLa cells with rabbit

polyclonal SLC35A4-MP antiserum validates SLC35A4-MP localization to mitochondria (SLC35A4-MP, green; MitoTracker (mitochondria), red; Hoechst (nuclei), blue). Scale bar: 10 μm. (E) Western blot analysis of total cell lysate and subcellular fractions prepared from HEK293T cells using a rabbit polyclonal SLC35A4-MP antiserum shows SLC35A4-MP localization to the mitochondria. (F) The mitochondrial fraction of HEK293T cells was treated with proteinase K under different conditions, showing that SLC35A4-MP is an inner mitochondrial membrane protein. Tom20 is the outer membrane marker, and TIMM50 is the inner membrane marker. Western blotting was performed using the rabbit polyclonal antiserum against SLC35A4-MP.

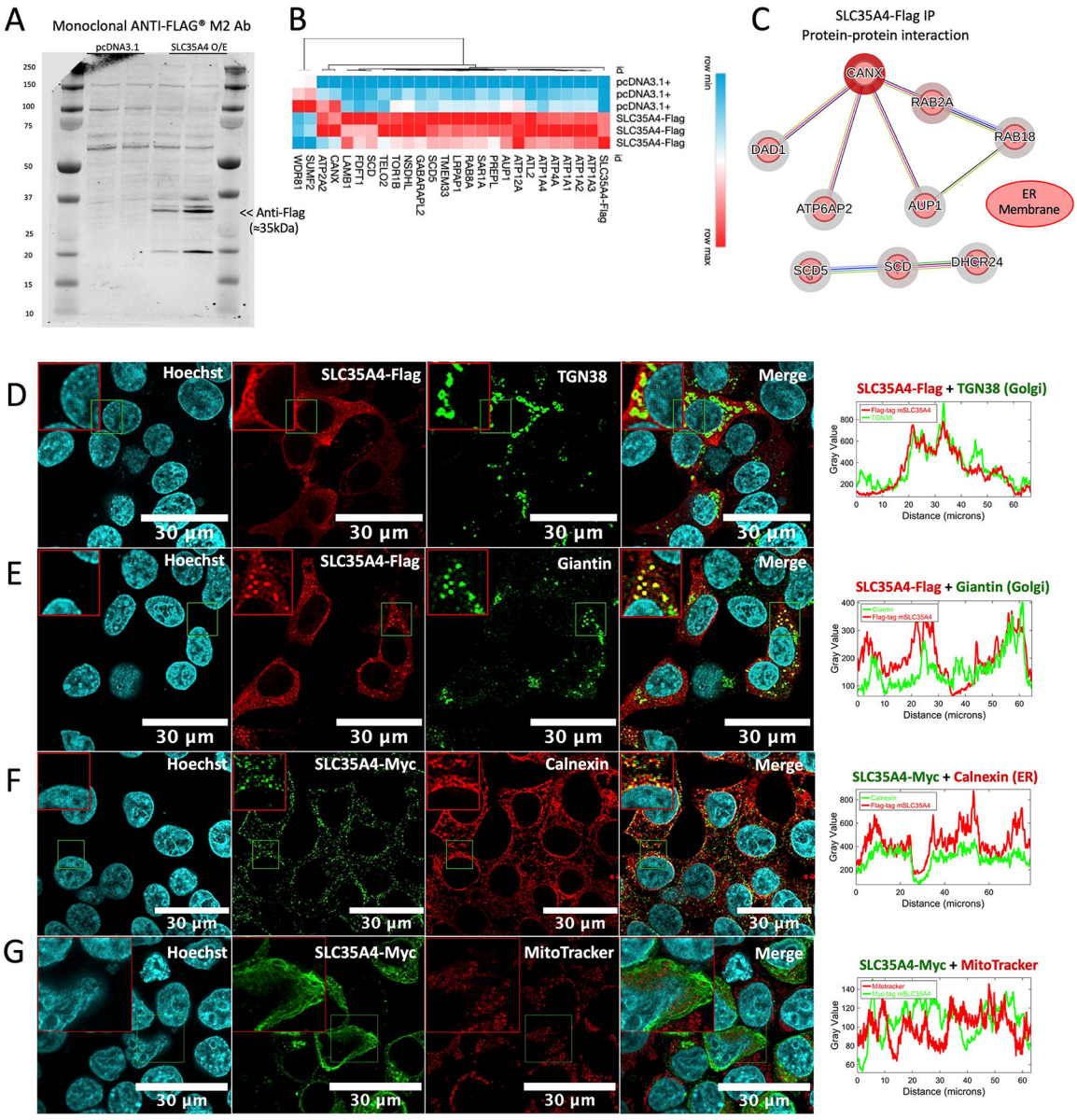

Fig. 3. Mass Spectrometry Analysis Shows No Evidence of Endogenous SLC35A4 Protein Expression.

(A) Anti-FLAG expression using Monoclonal ANTI-FLAG® M2 antibody (F3165) in cells overexpressing pSLC35A4 (Myc-DDK-tagged) was analyzed by Western blot and we identified (n=2 per group). (B) Heatmap showing proteins regulated by cells overexpressing SLC35A4-Flag detected by TMT-based proteomics (n=3 for control and FLAG groups). (C) STRING Protein-Network showing ER proteins enriched with immunoprecipitation of FLAG proteins sample (n=1 pool of the three FLAG samples for IP FLAG). (D and E) Colocalization of SLC35A4 (Myc-DDK-tagged) in the Golgi complex was analyzed in cells staining with Anti-Flag antibody in red, TGN38, and Giantin in green as a positive marker of the Golgi complex, and Hoechst stain for cell nuclei (blue). The colocalization of SLC35A4 (Myc-DDK-tagged) with Golgi compartments in cells was measured by fluorescence intensity analysis with ImageJ software. Scale bar, 30 µm. (F and

G) Colocalization of SLC35A4 (Myc-DDK-tagged) in the ER compartment was analyzed in cells staining with Anti-Myc-tag in green, Calnexin, and MitoTracker in red, and Hoechst stain for cell nuclei (blue). The colocalization of SLC35A4 (Myc-DDK-tagged) with ER and mitochondria compartments in cells was measured by fluorescence intensity analysis with ImageJ software. Scale bar, 30 μm. Values are the means ± SEM. *P < 0.05 versus pcDNA3.1+ (empty vector control group) (Mann-Whitney U test).

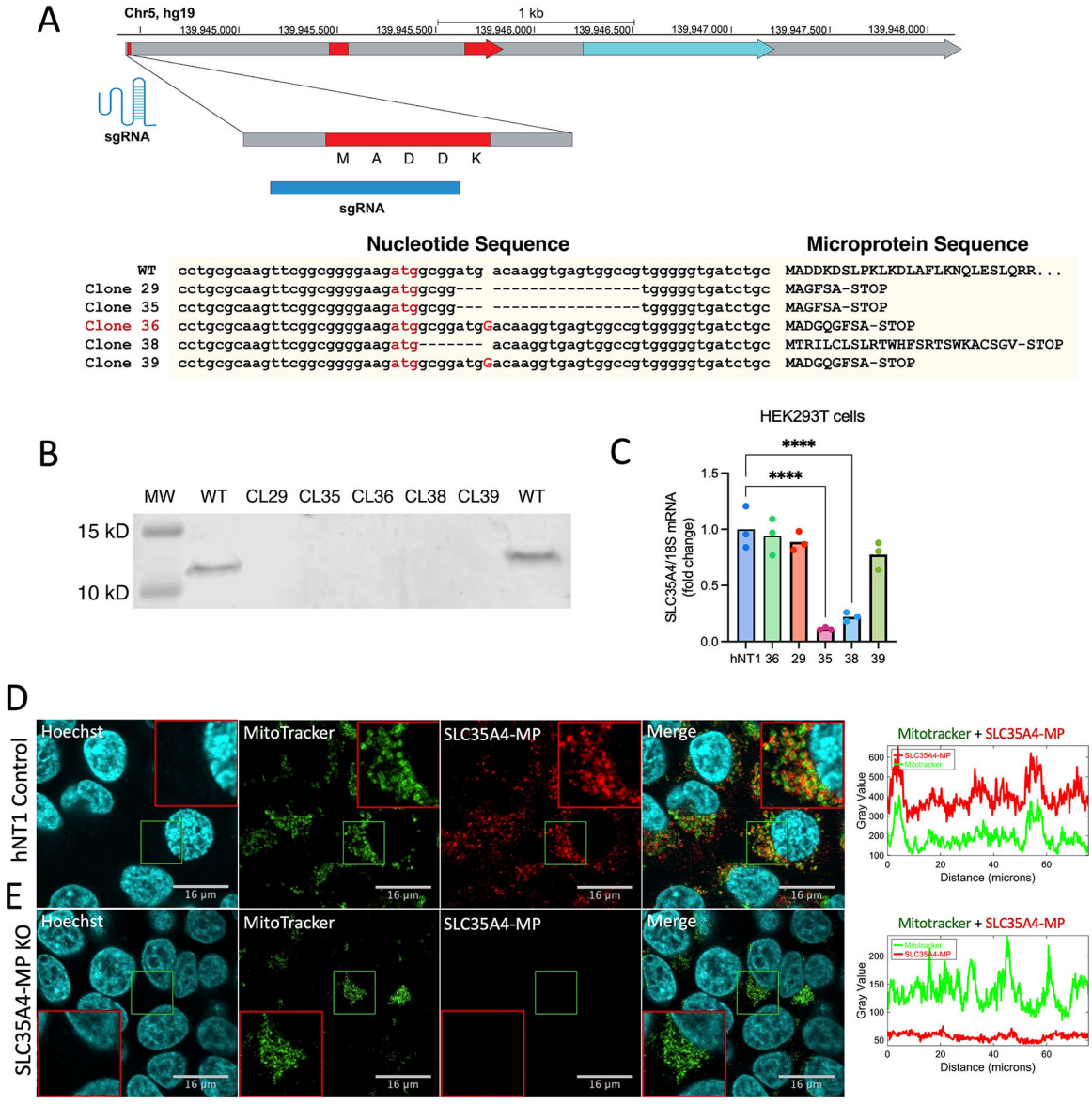

Fig. 4. Generating monoclonal cell lines targeting uORF of SLC35A4.

(A) The genotyping figure shows sequence alignment for the different clonal cell lines. (B) Validation of the rabbit polyclonal antiserum raised against the N-terminal 34 amino acids of SLC35A4-MP was performed through Western blot (WB) analysis. HEK293T cells were utilized, including a CRISPR-Cas9 control cell line (WT) and monoclonal SLC35A4 uORF knockout cell lines (KO29, KO35, KO36, KO38, and KO39). (C) qPCR analysis of SLC35A4 mRNA (n=3 per group). (D - E) Colocalization of SLC35A4-MP with MitoTracker Green, labels active mitochondria, was analyzed in HEK293T hNT1 and SLC35A4-MP KO (clone 36) cells staining with rabbit polyclonal SLC35A4-MP antiserum in red, MitoTracker Green in green, and Hoechst stain for cell nuclei (blue). The colocalization of SLC35A4-MP with mitochondria was measured by fluorescence intensity analysis with ImageJ software. Scale bar, 16 µm. Values are the means ± SEM. *P < 0.05 versus hNT1 (CRISPR control nontargeting sequence) (Mann-Whitney U test).

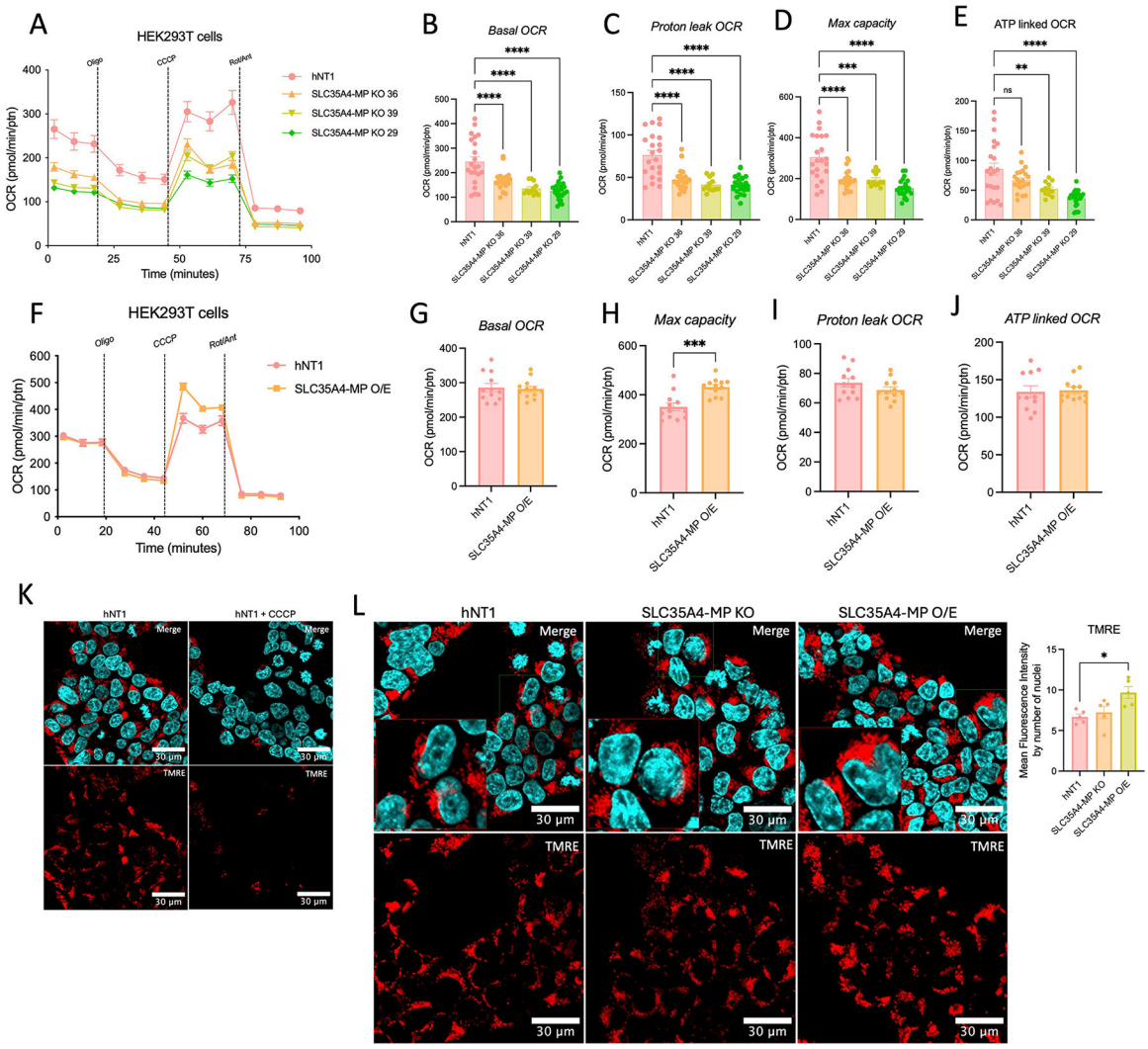

Fig. 5. Oxygen Consumption Rate Profiles of Clonal Cell Lines with SLC35A4-MP Knockout.

(A – E) Oxygen Consumption Rate Profiles were monitored in various clonal cell lines, including SLC35A4-MP knockout (KO) lines (36, 39, and 29), as well as the control cell line, hNT1, utilizing a Seahorse XFe96 analyzer (n=22 per group). (F – J) Oxygen Consumption Rate Profiles in HEK293T Cells Overexpressing SLC35A4-MP were analyzed using a Seahorse XFe96 analyzer. Metabolic inhibitors were introduced at various time points, as indicated (n=12 per group). (K and L) The mitochondrial membrane potential of HEK293T hNT1, SLC35A4-MP KO (clone 36), and SLC35A4-MP O/E cells was measured by TMRE (tetramethylrhodamine ethyl ester) staining. CCCP (a mitochondrial oxidative phosphorylation uncoupler) was used as system control. Representative confocal images are shown. Fluorescence intensity was analyzed using ImageJ software. Scale bar, 30 µm. Values are the means ± SEM of five replicates. N=5 per group. *P < 0.05 versus hNT1 (non-targeted CRISPR control), determined using the unpaired t-test. (Note: ptn = total protein)

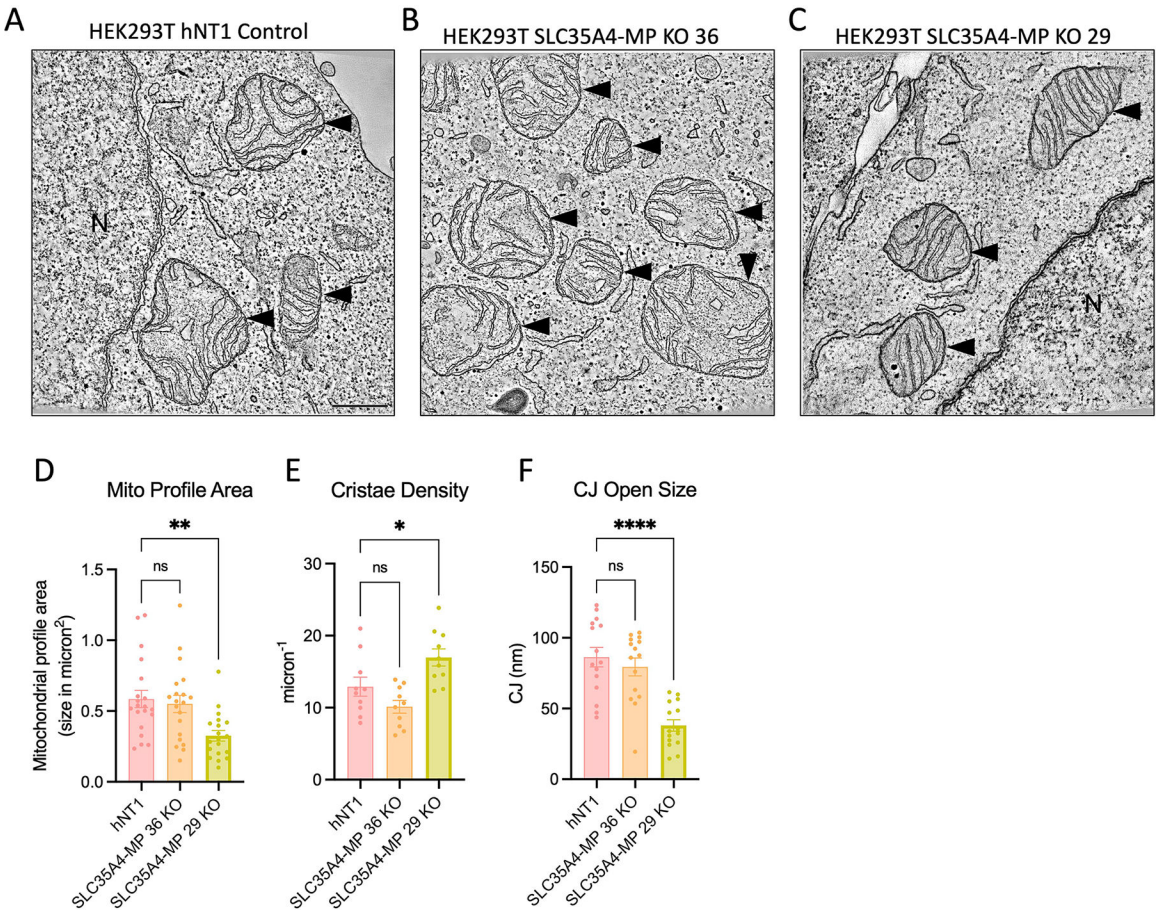

Fig. 6. CJ density and CJ per cristae in hNT1 and SLC35A4-MP KO HEK293T cells.

(A – C) Slices of thickness 1.6 nm through the middle of EM tomography volumes of mitochondria from hNT1 and SLC35A4-MP KO (clonal cell lines 36, and 29) HEK293T cells. Scale = 0.5 microns and applies to all panels. Quantitation of crista junctions (CJ) comparing hNT1 and SLC35A4-MP KO HEK293T clonal cell lines (36, and 29). CJ density is defined as the total number of crista junctions in a mitochondrion divided by the outer membrane surface area of that mitochondrion (N = 10), and the number of CJs per crista is defined as the total number of junctions divided by the total number of cristae in a mitochondrion (N = 15). Mito Profile Area is defined as the total mitochondria size (N=20). The arrows indicate mitochondria, and for the hNT and KO29 images, “N” = cell nucleus. Values are the means ± SEM. *P < 0.05 versus hNT1 (CRISPR control nontargeting sequence) (Mann-Whitney U test).