

## ARTICLE OPEN



## A naturally occurring variant of SHLP2 is a protective factor in Parkinson's disease

Su-Jeong Kim<sup>1</sup>, Brendan Miller<sup>1</sup>, Nicolas G. Hartel<sup>2</sup>, Ricardo Ramirez II<sup>1</sup>, Regina Gonzalez Braniff<sup>1</sup>, Naphada Leelaprachakul<sup>1,3</sup>, Amy Huang<sup>1</sup>, Yuzhu Wang<sup>1</sup>, Thalida Em Arpawong<sup>1</sup>, Eileen M. Crimmins<sup>1</sup>, Penglong Wang<sup>4</sup>, Xianbang Sun<sup>5</sup>, Chunyu Liu<sup>4,5,6</sup>, Daniel Levy<sup>5,6</sup>, Kelvin Yen<sup>1</sup>, Giselle M. Petzinger<sup>7</sup>, Nicholas A. Graham<sup>1,2,8</sup>, Michael W. Jakowec<sup>7,9</sup> and Pinchas Cohen<sup>1</sup>✉

© The Author(s) 2023

Mitochondrial DNA single nucleotide polymorphisms (mtSNPs) have been associated with a reduced risk of developing Parkinson's disease (PD), yet the underlying mechanisms remain elusive. In this study, we investigate the functional role of a PD-associated mtSNP that impacts the mitochondrial-derived peptide (MDP) Small Humanin-like Peptide 2 (SHLP2). We identify m.2158 T > C, a mtSNP associated with reduced PD risk, within the small open reading frame encoding SHLP2. This mtSNP results in an alternative form of SHLP2 (lysine 4 replaced with arginine; K4R). Using targeted mass spectrometry, we detect specific tryptic fragments of SHLP2 in neuronal cells and demonstrate its binding to mitochondrial complex 1. Notably, we observe that the K4R variant, associated with reduced PD risk, exhibits increased stability compared to WT SHLP2. Additionally, both WT and K4R SHLP2 show enhanced protection against mitochondrial dysfunction in in vitro experiments and confer protection against a PD-inducing toxin, a mitochondrial complex 1 inhibitor, in a mouse model. This study sheds light on the functional consequences of the m.2158 T > C mtSNP on SHLP2 and provides insights into the potential mechanisms by which this mtSNP may reduce the risk of PD.

*Molecular Psychiatry* (2024) 29:505–517; <https://doi.org/10.1038/s41380-023-02344-0>

## INTRODUCTION

Mitochondria are implicated in the pathogenesis of idiopathic Parkinson's disease (PD) [1]. A defect of mitochondrial complex I has been observed in the substantia nigra dopaminergic neurons of postmortem PD brains [2]. Recent study showed disruption of mitochondrial complex I induces progressive parkinsonism [2]. The complex I inhibitor MPTP (N-methyl-4-phenyl-1,2,3,6-tetrahydropyridine) causes parkinsonism and nigrostriatal dopaminergic degeneration in humans and animal models [3, 4]. Several monogenic forms of PD are caused by mutations in genes coding for mitochondrial function. For instance, the second most common cause of autosomal recessive early-onset PD can be attributed to mutations in the PINK1 gene, a mitochondrial serine/threonine kinase [1]. PINK1 plays a significant role in maintaining mitochondrial homeostasis, enhancing mitochondrial fission, and promoting mitophagy [1]. Conversely, the loss of PINK1 leads to a wide range of mitochondrial dysfunction [1]. Moreover, genetic studies of large cohorts of ethnically homogeneous subjects revealed that mitochondrial DNA (mtDNA) polymorphisms play a role in the risk of developing PD [5–8]. Although a strong association has been found by population genetics, the functional effects of the mtDNA polymorphisms are underexplored.

MDPs are bioactive microproteins that are encoded from small open reading frames in mitochondrial DNA. Since the first MDP

named humanin was discovered in 2001, eight more microproteins have been identified that play essential roles in pathophysiology of age-related diseases [9]. The levels of some MDPs decline with age in mice and humans, and the administration of MDPs can exert beneficial effects in in vitro and in vivo models of age-related disease, including Alzheimer's disease, cardiovascular disease, and type 2 diabetes (T2D) [10, 11]. One of the MDPs, a small humanin-like peptide 2 (SHLP2), is encoded from the non-coding RNA ASncmtRNA-1/2 in the mtDNA and plays an essential role in various cellular processes [12]. SHLP2 improves mitochondrial metabolism by increasing respiration and biogenesis, reducing reactive oxygen species, and decreasing mtDNA oxidation [12]. SHLP2 has also been suggested to be a possible treatment for age-related macular degeneration (AMD). Work from the Kenney lab has found that SHLP2 administration was able to prevent a number of negative effects in an in vitro AMD model [13]. The misfolding and subsequent aggregation of Islet amyloid polypeptide (IAPP) induces beta cell apoptosis, oxidative damage, and mitochondrial dysfunction, playing critical roles in the pathogenesis of type 2 diabetes (T2D). SHLP2 acts as a chaperone and directly binds to the misfolded, seeding-capable IAPP species and blocks amyloid seeding [14]. The chaperone-like activity can link the anti-apoptotic roles and beneficial metabolic effect of SHLP2 against IAPP aggregation, suggesting that SHLP2 has

<sup>1</sup>The Leonard Davis School of Gerontology, University of Southern California, Los Angeles, CA, USA. <sup>2</sup>Mork Family Department of Chemical Engineering and Materials Science, University of Southern California, Los Angeles, CA, USA. <sup>3</sup>Environmental Toxicology Program, Chulabhorn Graduate Institute, Bangkok 10210, Thailand. <sup>4</sup>The Population Sciences Branch, National Heart, Lung, and Blood Institute, National Institutes of Health, Bethesda, MD, USA. <sup>5</sup>Department of Biostatistics, Boston University School of Public Health, Boston, MA, USA. <sup>6</sup>Boston University's and National Heart, Lung, and Blood Institute's Framingham Heart Study, Framingham, MA, USA. <sup>7</sup>Department of Neurology, University of Southern California, Los Angeles, CA, USA. <sup>8</sup>Norris Comprehensive Cancer Center, University of Southern California, Los Angeles, CA, USA. <sup>9</sup>Department of Biokinesiology and Physical Therapy, The George and MaryLou Boone Center for Parkinson's Disease Research, University of Southern California, Los Angeles, CA, USA. ✉email: hassy@usc.edu

Received: 9 July 2022 Revised: 20 November 2023 Accepted: 27 November 2023

Published online: 3 January 2024

potential as a T2D therapeutic. Circulating levels of SHLP2 have also been correlated with android and liver fat and inversely correlated with prostate cancer risk [15, 16].

Recent studies show that mitochondrial DNA single nucleotide polymorphisms (mtSNPs) within coding regions of MDPs are associated with age-related deficits and could help discover novel microproteins. For example, mtSNP m.2706 A > G in humanin predicts lower circulating humanin levels and worse cognitive decline [17]. Administration of humanin in old mice prevents age-related cognitive decline [17]. The m.12372 G > A is associated with Alzheimer's disease (AD) and revealed a novel MDP called SHMOOSE [11]. Additionally, m.1382 A > C – a mtSNP that changes the 14th amino acid of the wild-type MDP MOTS-c from the amino acid lysine (K) to glutamine (Q) – predisposes individuals of East Asian descent to Type 2 Diabetes [18]. Importantly, K14Q MOTS-c is less effective as a metabolic regulator in cell and animal models and is associated with suppressed insulin-sensitizing effects and higher body fat levels [18].

The mtSNP m.2158 T > C is associated with reduced risk of PD [6]. However, the underlying biological mechanisms responsible for this effect remains a gap in our knowledge. Here, we present a series of experiments that explore the effects of the m.2158 T > C allelic form of SHLP2 (termed K4R SHLP2) on mitochondrial function. Elucidating mechanisms by which SHLP2 and the K4R variant impact mitochondrial function may guide novel therapeutic modalities focused on restoring mitochondrial function in PD.

## MATERIALS AND METHODS

### Human cohorts

#### Health Retirement Study (HRS)

**Participants:** Data are from the Health and Retirement Study (HRS), a nationally representative sample of older Americans over age 50 [19, 20] in the contiguous United States. The present analysis was limited to participants who self-reported their race as white/Caucasian, verified by principal components analysis of ancestry markers, in order to assess effects of the SHLP2 minor allele only found in European ancestry groups. The analytical sample for the HRS included individuals who had available genetic data, at least one measure of total cholesterol data (assessed in 2006, 2010, 2012, or 2014) and relevant covariate data ( $N = 9,774$ ), with a sub-sample available for the analysis with LDL (assessed in 2016;  $N = 4,297$ ). The HRS data collection is supported by: U01 AG009740 to the University of Michigan.

**SHLP2 Mitochondrial Single Nucleotide Polymorphism (mtSNP):** For HRS, genotype data were accessed from the National Center for Biotechnology Information Genotypes and Phenotypes Database (dbGaP [21]). Genotyping was conducted on over 15,000 individuals using either the Illumina HumanOmni2.5-4v1 (2006 and 2008) and HumanOmni2.5-8v1 (2010) arrays and was performed by the NIH Center for Inherited Disease Research (CIDR). Standard quality control procedures were implemented by the University of Washington Genetic Coordinating Center [22]. Further detail is provided in the HRS documentation [23]. The SHLP2 allele coding (binary coded as 0 or 1 for a mitochondrial SNP) was extracted using PLINK 1.9 [24, 25]. **Cholesterol Biomarkers:** The HRS collected blood-based biomarkers including total cholesterol and high-density lipoprotein (HDL) levels on half of the sample in 2006, and the other half in 2008, with additional individuals captured in the 2010 or 2012 data collection waves. Detailed information on collection and assay are provided elsewhere [26, 27]. Total cholesterol values were adjusted for medication use as done previously in the FHS study. For individuals reporting taking medications to lower their cholesterol, levels were divided by .80 to produce an adjusted total cholesterol value [28]. In 2016, venous blood samples were collected on a subsample in HRS to assay and calculate low density lipoprotein (LDL) levels with detailed information provided previously [29]. **Covariates:** In HRS, covariates included age at biomarker assessment, gender (0=female, 1=male), and five principal components to reduce type 1 error due to differences in underlying population substructure [30, 31]. Detailed descriptions of the processes employed for running principal components analysis, including SNP selection, are provided by HRS [23], and follow gold standard methods [32]. Models constructed for HDL and LDL were

adjusted for medication use (0=no, 1=yes), if respondents reported using medication to lower their cholesterol.

**Statistical analysis:** In HRS, multivariable linear regression models were run to test for the association between SHLP2 and dependent variables of total cholesterol (medication adjusted value), HDL and LDL, adjusting for age, gender, principal components, and medication use (for HDL and LDL) using SAS 9.4.

#### Cardiovascular Health Study (CHS) and Framingham Heart Study (FHS)

**Participants:** The Cardiovascular Health Study (CHS) ( $n = 2,772$ ): The Cardiovascular Health Study (CHS) originated in 1988 to recruit participants from four U.S. communities [33]. The CHS has been followed since then to study coronary heart disease and stroke in 5,888 elderly adults aged 65 years and older. The first exam began in June 1989. A second comprehensive exam began 3 years after the first exam. All study participants provided written informed consent for genetic studies. A total of  $n = 2,772$  CHS participants of white/Caucasian (mean age 74 and 57% women) with whole genome sequencing were included in this study. Cohort acknowledgement/support: The CHS (phs001368.v1.p1) was supported by contracts 75N92021D00006, HHSN268201200036C, HHSN268200800007C, HHSN268201800001C, N01HC55222, N01HC85079, N01HC85080, N01HC85081, N01HC85082, N01HC85083, N01HC85086, and grants U01HL080295, R01HL105756, and U01HL130114 from the NHLBI, with additional contribution from the National Institute of Neurological Disorders and Stroke (NINDS). Additional support was provided by R01AG023629 from the National Institute on Aging (NIA). A full list of principal CHS investigators and institutions can be found at CHS-NHLBI.org. Sequencing was supported and conducted in collaboration with Baylor University (HHSN268201600033I, 3U54HG003273-12S2, HHSN268201500015C) contracts from NHLBI.

**The Framingham Heart Study (FHS) ( $n = 3,621$ ):** The FHS is a single-site, community-based, prospective study that was initiated in 1948 to investigate the risk factors for CVD [34]. The second generation [35] was recruited in 1971 and the third generation [36] was recruited between 2002 and 2005. The first generation has been examined every 2 years. The second generation has been examined every 4–8 years. The third generation has had three examinations. A small number of spouse individuals of the second generation was examined at the same time when the third generation had their first examination. A total of 4,196 FHS participants were whole genome sequenced by TOPMed; of those, 376 were the first generation, 2218 were the second generation and 95 were spouses of the second generation participants; and 1507 were the third generation participants. All study participants provided written informed consent for genetic studies. This study included 3,621 FHS participants. The WGS for FHS (phs000974) was performed at the Broad Institute of MIT and Harvard (3R01HL092577-06S1 and 3U54HG003067-12S2). The FHS acknowledges the support of contracts NO1-HC-25195, HHSN268201500001I and 75N92019D00031 from the National Heart, Lung and Blood Institute and grant supplement R01 HL092577-06S1 for this research. We also acknowledge the dedication of the FHS study participants without whom this research would not be possible. Dr. Vasan is supported in part by the Evans Medical Foundation and the Jay and Louis Coffman Endowment from the Department of Medicine, Boston University School of Medicine. X.L., S.S., C.L.S. and C.L. are also supported by R01AG059727. C.L.S. and S.S. are also supported by AG052409, AG054076 and AG059421.

**Cholesterol biomarkers:** At each of the health exams, the FHS collected blood-based biomarkers including total cholesterol, high-density lipoprotein (HDL), and triglyceride. We matched these lipid measurements to the time when blood was drawn for whole genome sequencing LDL (mg/dL) was calculated as  $(TC - HDL - TRIG/5)$  in individuals with  $TRIG < 400$  mg/dL TC values [37].

**Covariates:** Covariates included age at whole genome sequencing, gender (0=female, 1=male), and five principal components to reduce type 1 error due to differences in underlying population substructure [30, 31]. In TOPMed, principle components calculated PCA using PC-Air [38]. Models constructed for HDL and LDL were adjusted for medication use (0=no, 1=yes), if respondents reported using medication to lower their cholesterol.

**mtDNA homoplasmy derivation in FHS and CHS:** Whole blood derived DNA was used for WGS from TOPMed sequencing centers. In

analyzing sequencing data, the coverage was defined as the number of reads that were mapped to a given nucleotide in the reconstructed sequence. The average coverage was ~39x across samples in TOPMed. The program *fastMitoCalc* of the software package *mitoAnalyzer* [39] was used to identify mtDNA homoplasmy based on derived alternative allele fractions (AAFs) after comparing sequencing reads to rCRS [40]. We set an mtDNA locus as missing if the coverage was <250-fold [39]. We compared homoplasmic alleles called by previous genotyping arrays to the ones derived from TOPMed whole genome sequencing in the same individuals [41]. Individuals with >two inconsistent homoplasmic alleles (out of ~200 mtDNA variants genotyped by the arrays) were red-flagged and discrepancies examined. We also compared the homoplasmic variants within maternal lineage members. We also removed several sites listed in 'blacklisted sites' (301,302,310, 316, 3107, and 16182 mtDNA loci) recommended by GATK (<https://console.cloud.google.com/storage/browser/gatk-best-practices/mitochondria-pipeline/>). We determined that the >97% AAF was used to determine homoplasmy based on the calling results from 3 samples that each had 4 repeated whole genome sequencing data.

**Statistical analysis:** In FHS and CHS, multivariable linear regression models were performed to test for the association between SHLP2 and dependent variables of total cholesterol (medication adjusted value), HDL and LDL, adjusting for age, gender, five principal components, and medication use (for HDL and LDL) using R software version 4.0.0.

**Mass spectrometry of SHLP2 identification.** SH-SY5Y cells were differentiated using 10  $\mu$ M of retinoic acid, as previously described [42]. Approximately  $8 \times 10^7$  differentiated cells were washed once with ice cold PBS and immediately lysed using 50 mM HCl, 2 mM TCEP, 0.05% Triton X-100 (lysis buffer) at room temperature for 15 minutes. This lysis buffer has previously yielded high yield of microproteins from cell extracts [43]. After the 15-minute lysis duration, the sample was centrifuged at 25,000  $\times$  g for 30 minutes at 4°C. Cleared lysate was then passed through a 5  $\mu$ m strainer (pluriSelect Life Science, Leipzig, Germany). A BCA protein assay (Thermo Scientific, Waltham, MA) was used to quantify protein concentration. Synthetic light (MGVKFFTLSTRFFPSVQRAVPLWTNS) and heavy (MGVKFFTLSTRFFPSVQ[Arg(13C6 15N4)]AVPLWTNS) versions sequence of SHLP2 were combined and reduced with dithiothreitol, alkylated with iodoacetamide, digested with trypsin and STAGE tip desalted on a C8 core. Digests were loaded onto a nanoscale UHPLC system (EASY-nLC1200, Thermo Scientific) connected to a Q Exactive Plus hybrid quadrupole-Orbitrap mass spectrometer equipped with a nanoelectrospray source (Thermo Scientific). Peptides were separated by a reversed-phase analytical column (PepMap RSLC C18, 2  $\mu$ m, 100 Å, 75  $\mu$ m  $\times$  25 cm) (Thermo Scientific). Flow was set to 200 nL/min starting with 3% buffer B (0.1% formic acid, 80% acetonitrile) to 38%B in 110 minutes, then washed with 75%B for 1 minute and then 85%B for 10 minutes and held at 85%B for 9 minutes. Column temperature was set to 55°C. Spray voltage was set to 2000V and all experiments were run in positive mode. A series of injections ranging from 700 pg – 500 ng of protein were performed on a data-dependent acquisition (DDA) method and data imported to Skyline where a parallel reaction monitoring (PRM) inclusion list was generated for the most common peptides identified from both the light and heavy SHLP2. The DDA method full scan range was from 400–900 m/z with a 70,000 resolution and maximum injection time of 120 ms with 1e6 AGC target, and 17,500 resolution for ddMS2 scans with a maximum injection time of 120 ms and 1e5 AGC target. Isolation width was set to 1.0 m/z with a normalized collision energy of 26. Dynamic exclusion was enabled and set to 20 seconds. The PRM method was run at 35,000 resolution with a 700 ms injection time and 5e5 AGC target with the same isolation width and collision energy as the DDA method. Lysate from SH-SY5Y cells were enriched by C8 column enrichment, concentrated by methanol/ chloroform precipitation, reduced, alkylated, digested with trypsin, desalted with C8 STAGE tip and run in PRM mode with the inclusion list. A parallel procedure was done with the heavy labeled version of SHLP2 spiked into SH-SY5Y lysate. RAW files were searched both in Skyline and Proteome Discoverer 2.2 using SEQUEST and MS AMANDA 2.0 using a mitochondrial protein FASTA file containing SHLP2 sequence. Cleavage was set to trypsin allowing for up to two missed cleavages, with oxidation of methionine (+15.995 Da), heavy arginine (+10.008 Da), and acetylation of protein N-terminus (+42.011 Da) set as variable modifications. Carbamidomethylation of cysteine (+57.021 Da) was set as a static modification. Precursor mass tolerance was set to 10 ppm and fragment ion tolerance was set to 0.02 Da. Validation was performed by Fixed Value PSM Validator node in

Proteome Discoverer with maximum delta Cn set to 0.05. Identified peptides were further confirmed by checking exact m/z values and retention times to those observed in the synthetic protein experiments.

**Reagents, plasmids, antibodies.** SHLP2 (small humanin-like peptide 2) peptides were synthesized by Genscript (Piscataway, NJ, USA). Peptides were initially dissolved in Milli-Q water. Cycloheximide (Sigma, St. Louis, MO, USA) was used for pulse chase experiment in culture. pcDNA3.1-WT SHLP2-EGFP and pcDNA3.1-K4R SHLP2-EGFP were cloned in Genscript (Piscataway, NJ, USA). Transient transfection was performed with Lipofectamine 3000 as recommended by the supplier (ThermoFisher Scientific). After transfection, the cells were incubated for 5 h, switched into a complete culture medium, and further incubated for an additional 36 h. The following antibodies were used in this study: GFP (1:1000), Lamin B1 (1:1000), S6K (1:1000), anti-rabbit IgG, HRP-linked antibody (1:30000), and anti-mouse IgG, HRP-linked antibody (1:30000). (all Cell Signaling Technology), ACTIN (1:2000) and GRSF-1 (1:1000; sigma), Tom20 (1:100; Santa Cruz Biotechnology). Customized rabbit anti-SHLP2 antibody was produced by Harlan Laboratories (Indianapolis, IN, USA).

**Cell culture and treatment.** HEK293 were purchased from ATCC (Manassas, VA, USA) and HeLa expressing HA-EGFP-OMP25 (HA-MITO) or myc-EGFP-OMP25 (CON-MITO) cells were gifts from Dr. Sabatini [44]. The HEK293 and HeLa cells were cultured in DMEM + 10% FBS (FBS; ThermoFisher Scientific) at 37°C in 5% CO<sub>2</sub>. Cycloheximide were treated into HEK293 cells for the indicated time. Wild type and TFAM heterozygous mouse embryonic fibroblasts cells were Cultured in DMEM + 15% FBS at 37°C in 5% CO<sub>2</sub>.

**Western blot analysis.** Cells were lysed with RIPA Lysis and Extraction Buffer (ThermoFisher Scientific) plus the Halt protease & phosphatase inhibitor cocktail (ThermoFisher Scientific). The lysates were incubated on ice for 10 min then homogenized using a sonicator, and the supernatant was collected by centrifugation at 15,000  $\times$  g for 15 min at 4°C. Protein content in the cellular lysates was quantified using the Pierce™ BCA Protein Assay Kit (ThermoFisher Scientific). Predetermined amounts of proteins (10–30  $\mu$ g) were separated on 8–16% SDS-PAGE gels and blotted onto PVDF membranes (Biorad, Hercules, CA, USA). Membranes were incubated with primary antibody at 4°C overnight according to the manufacturer's instructions. After several washes with Tris-buffered saline containing 0.1% Tween-20, membranes were incubated at room temperature for 1 hr with the appropriate HRP-conjugated secondary antibody. Clarity™ Western ECL substrate (Biorad) was used for detecting specific bands. Membranes were imaged on a Bio-Rad ChemiDoc XRS<sup>+</sup> imager. If necessary, relative intensities of the bands in each condition were measured using Image J, a free software program provided by National Institute of Health (Bethesda, Maryland, USA).

**Mitochondria isolation using HeLa expressing HA-EGFP-OMP25 (HA-MITO).** Mitochondria isolation from HeLa cells expressing HA-EGFP-OMP25 (HA-MITO) or myc-EGFP-OMP25 (CON-MITO) was performed according to the published article [44]. Briefly, cells were lysed in KPBS buffer (136 mM KCl, 10 mM KH<sub>2</sub>PO<sub>4</sub>, pH 7.25) and homogenized with 25 strokes of a 2 ml homogenizer. The homogenate was spun down at 1000 g for 2 min. The supernatant was incubated with anti-HA magnetic beads on an end-over-end rotator for 3.5 min. The beads were washed with KPBS buffer, incubated for 10 min with lysis buffer (1% triton X-100 lysis buffer with protease inhibitor cocktail), and spun down at 17000 g for 10 min. The supernatants were used for western blot.

**Mitochondria and cytoplasmic fractionation.** Mitochondrial and cytoplasmic fractionation from HEK293 cells was performed using the Qproteome Mitochondrial Isolation kit (Qiagen, Valencia, CA, USA) according to the manufacturer's instruction. Briefly, cells were resuspended in Lysis Buffer and incubated for 10 min. Cell lysates were centrifuged at 1,000  $\times$  g for 10 min at 4°C. The supernatant was collected for the cytosolic fraction. The pellet was resuspended in Disruption Buffer. After homogenization, resuspension was centrifuged at 1,000  $\times$  g for 10 min at 4°C. The supernatant was collected and centrifuged at 6,000  $\times$  g for 10 min. The pellet contains mitochondria.

**Sodium carbonate extraction.** Mitochondrial extraction from HEK293 cells was performed using the Qproteome Mitochondrial Isolation kit (Qiagen) according to the manufacturer's instruction. Mitochondria pellets isolated



from HEK293 cells were resuspended in buffer (320 mM Sucrose, 1 mM EDTA, 10 mM Tris-Cl pH 7.4) containing either 1% Triton X or 0.1 M Na<sub>2</sub>CO<sub>3</sub> pH 11. The resuspensions were incubated at 4 °C for 1 h with gentle nutation. Supernatant and insoluble pellet were separated by centrifugation at 20,000 × g for 30 min, and the supernatant was gently removed without disturbing the pellet. Equal fractions of supernatant and pellet were analyzed by SDS-PAGE immunoblotting.

**Immunocytochemistry.** HEK293 cells cultured on coverslips were transiently transfected with WT or K4R SHLP2-EGFP for 36 hr and were cultured on coverslips and then fixed with 4% paraformaldehyde for 10 min at room temperature. After fixation, the cells were permeabilized with 0.2% Triton X-100 in phosphate-buffered saline (PBS) for 10 minutes at room temperature and were blocked in PBS containing 0.2% Triton X-100 and 1% bovine serum albumin (BSA) for 1 hour at room temperature. Cells were then incubated with mouse anti-Tom20 antibody (1:100; Santa Cruz Biotechnology) in PBS containing 0.2% Triton X-100 and 1% BSA at 4 °C overnight. After three washes with PBS, the cells were further incubated with Alexa Fluor 568-conjugated donkey anti-mouse IgG (1:200; Invitrogen) in PBS containing 0.2% Triton X-100 and 1% BSA for 1 hour at room temperature in dark. Nuclei were stained for 5 minutes at room temperature in PBS containing Hoechst 33258 (2 mg/ml; Invitrogen). Coverslips were mounted with ProLong Gold antifade reagent (Invitrogen). Images were acquired with a Keyence microscope (Keyence corporation of America, Itasca, IL). For immunostaining of COX IV antibody, fixed and permeabilized cells were incubated with anti-COX IV antibody (1:200; abcam) and Alexa Fluor 448-conjugated donkey anti-rabbit IgG (1:200; Invitrogen).

**RNA extraction and qRT-PCR.** Total RNA was extracted from cells and tissues using the Direct-zol RNA MiniPrep Plus (ZYMO RESEARCH, Irvine, CA, USA). The protocol was performed according to manufacturer's instructions and RNA was measured at 260 nm with the Nanodrop system (ThermoFisher Scientific, Wilmington, DE, USA). RNA purity and quality were evaluated by 260/280 and 260/230 ratios. 1 mg of RNA was used for reverse transcription by using SuperScript IV Reverse Transcriptase and oligo dT 20 primers (ThermoFisher Scientific) according to the manufacturer's instruction. For quantitative real-time PCR (CFX system, Biorad) SsoAdvanced Universal SYBR green supermix (Biorad) was used to amplify cDNA and quantify the relative gene expression analysis. The 2<sup>-ΔΔCT</sup> method is used for relative gene expression analysis. This method shows the fold increase (or decrease) of the target gene in the test sample relative to the calibrator sample and is normalized to the expression of a reference gene. First, we normalized the C<sub>T</sub> of the target gene to that of the reference gene for both the test and calibrator sample (ΔC<sub>T</sub>). Secondly, we normalized the ΔC<sub>T</sub> of the test sample to the ΔC<sub>T</sub> of the calibrator sample (ΔΔC<sub>T</sub>). Finally, we calculated the expression ratio, 2<sup>-ΔΔCT</sup>. RPS13 was used as reference gene. The primers used for amplification are: GFP Forward primer (5' GAACCGCATCGAGCTGAA 3') and Reverse primer (5' TGCTTGTCGGCCATGATATAG 3') and RPS13 Forward primer (5'CATGGCTCGCTCGTGAC3'), and Reverse primer (5'CAGTTCAGTATGTTCCGCTCC3').

**Pulse Chase experiment.** Cycloheximide, an inhibitor of protein biosynthesis, was used to examine protein stability in cells. HEK293 cells were seeded in 6-well plate. WT and K4R SHLP2 were transfected into HEK293 cells. After 24 hr, 100 μg/ml cycloheximide was treated for the indicated time. The protein quantity from western blot was examined with Image J.

**DNA extraction and mitochondrial DNA copy number measurement.** Genomic DNA from 300,000 cells was extracted with the DNeasy Blood & Tissue Kits (Qiagen). The mitochondrial copy number was estimated by real-time PCR (CFX Connect Real-Time System, Biorad) using two mtDNA targets (ND1, CYB) and two nuclear DNA targets (ApoB, B2M) (IDT, CA, USA). Real-time PCR was performed by using SsoAdvancedTM Universal SYBR® Green Supermix, following the protocol provided by the manufacturer. The ratio of mtDNA to nuclear DNA was calculated by averaging the differences in amplification efficiency between ND1 and ApoB (ΔCt) and between CYB and B2M (ΔCt). The primers used for amplification are: ND1 Forward primer (5'-gccgtagcccaacaatttc-3') and Reverse primer (5'-caggctggcagaagtaatacata-3'); CYB Forward primer (5'-aggagaccagacaactaca-3') and Reverse primer (5'-tgagcgctagaatggcgtatg-3'); ApoB Forward primer (5'-cgtgggctccagcattcta-3') and Reverse primer (5'-tcaccagtcattctgcctttg-3'); B2M Forward primer (5'-atgggaagccgaacatactg-3') and Reverse primer (5'-cagctcagtggtgggtgaat-3').

**MTT assay.** The viability of cells after SHLP2 treatment were determined using the tetrazolium dye MTT [3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide] assay. All the treatments were done using 5,000 cells in 96 well plate. The purple insoluble formazan was read on a microplate reader at a wavelength of 570 nm.

**Biotin-Phenol labeling in live cells and western blot of biotin-phenol labeling.** Biotin-phenol labeling in live cells was performed according to previously published protocol [45, 46]. Briefly, the plasmid harboring SHLP2-APEX fusion proteins (SHLP2-APEX-myc) or the APEX control (Flag-APEX-myc) were transiently transfected into HEK293 cells using Lipofectamine 3000. Twenty-four hours post-transfection, the cell culture medium was changed to fresh growth medium containing 500 μM biotin-tyramide (CDX-B0270, Adipogen). After incubation at 37 °C for 30 min, H<sub>2</sub>O<sub>2</sub> was added to each plate at a final concentration of 1 mM, and the plates were gently agitated for 1 min. Cells were then washed three times with a quenching solution [5 mM Trolox, 10 mM sodium azide, and 10 mM sodium ascorbate in PBS], and the pellet was collected by centrifugation at 1000 g for 5 min. Cell pellets were lysed on ice for 20 min in RIPA buffer supplemented with a protease/phosphatase inhibitor and 1 mM phenylmethanesulfonyl fluoride (PMSF) followed by centrifugation at 20,000 g for 20 min at 4 °C to remove cell debris. Cell lysates were added to prewashed streptavidin agarose resin, rotated at 4 °C for 4 hr, and then washed three times. Bound proteins were eluted with 2X SDS loading buffer and analyzed by Western blotting.

**SHLP2 peptide column.** Isolated mitochondrial protein was prepared from HEK293 cells. The cells were scraped and homogenized in mitochondrial isolation buffer (20 mM potassium HEPES, pH7.4, 75 mM Sucrose, 10 mM EDTA, 215 mM Mannitol), followed by differential centrifugation using 700 g for 10 min and 10,500 g for 10 min, respectively. The mitochondrial pellets were lysed using 0.1% NP-40 lysis buffer (50 mM Tris-chloride, pH 8.0, 150 mM NaCl, 0.1% Nonidet P-40 Substitute 11332473001, Roche) with Halt™ Protease Inhibitor Cocktail (ThermoFisher). The mitochondrial lysate concentrations were measured using BCA method (ThermoFisher). The scrambled peptide and SHLP2 peptide were pre-immobilized on agarose beads from CarboxyLink™ kit (ThermoFisher). 1.6–1.9 mg of mitochondrial protein extract were incubated in the equilibrated column at 4 °C overnight on a rotator. The columns were washed, then the bound proteins were eluted by Elution Buffer according to the manufacturer's protocol. The 280 nm absorbance of each elution was measured to observe protein concentrations. The eluted proteins were concentrated using Amicon® Ultra-15 Centrifugal Filter Devices (Millipore) before being analyzed by mass spectrometry.

**Biotin-Phenol labeling mass spectrometry and SHLP2 peptide column mass spectrometry.** Biotin-Phenol labeling mass spectrometry- Digestion and Desalting: Proteins bound to streptavidin beads were reduced and alkylated via sequential 20 min incubations of 5 mM TCEP and 10 mM iodoacetamide at room temperature in the dark while being mixed at 1200 rpm in an Eppendorf thermomixer. Proteins were then digested by the addition of 0.1 μg Lys-C (Promega) and 0.8 μg Trypsin (Thermo Scientific, 90057) while shaking 37 °C overnight. The digested samples were quenched by addition of formic acid to a final concentration of 5% (v/v). Each sample was desalted through C18 tips (Thermo Scientific, 87784) and then resuspended in 5% formic acid before being analyzed by LC-MS/MS.

**SHLP2 peptide column mass spectrometry-Digestion and Desalting:** SHLP2 peptide column elution samples were mixed with same volume of digestion buffer (8 M Urea, 0.1 M Tris-HCl pH 8.5), then each sample was reduced and alkylated via sequential 20-minute incubations with 5 mM TCEP and 10 mM iodoacetamide at room temperature in the dark while being mixed at 1200 rpm in an Eppendorf thermomixer. 18 μl of carboxylate-modified magnetic beads (CMMB and also widely known as SP3) was added to each sample. Ethanol was added to a concentration of 50% to induce protein binding to CMMB. CMMB were washed 3 times with 80% ethanol and then resuspended with 50 μl 50 mM TEAB. The protein was digested overnight with 0.1 μg LysC (Promega) and 0.8 μg trypsin (Thermo Scientific, 90057) at 37 °C. Following digestion, 1 ml of 100% acetonitrile was added to each sample to increase the final acetonitrile concentration to over 95% to induce peptide binding to CMMB. CMMB were then washed 3 times with 100% acetonitrile and the peptide was eluted with 50 μl of 2% DMSO. Eluted peptide samples were dried by

vacuum centrifugation and reconstituted in 5% formic acid before analysis by LC-MS/MS.

**LC-MS Acquisition and analysis:** Peptide samples were separated on a 75  $\mu$ M ID, 25 cm C18 column packed with 1.9  $\mu$ M C18 particles (Dr. Maisch GmbH) using a 140-minute gradient of increasing acetonitrile concentration and injected into a Thermo Orbitrap-Fusion Lumos Tribrid mass spectrometer. MS/MS spectra were acquired using Data Dependent Acquisition (DDA) mode. MS/MS database searching was performed using MaxQuant (1.6.10.43) against the human reference proteome from EMBL (UP000005640\_9606 HUMAN Homo sapiens, 20874 entries). Statistical analysis of MaxQuant label-free quantitation data was performed with the artMS Bioconductor package which performs the relative quantification of protein abundance using the MSstats Bioconductor package (default parameters). The abundance of proteins missing from one condition but found in more than 2 biological replicates of the other condition for any given comparison were estimated by imputing intensity values from the lowest observed MS1-intensity across samples and *p* values were randomly assigned to those between 0.05 and 0.01 for illustration purposes.

**Co-expression network analysis.** The SHLP2 transcript was correlated to all gene transcripts using data from iPSCs-derived dopaminergic neurons from the Cedars-Sinai Board of Governors Regenerative Medicine Institute. Briefly, counts from RNASeq data were produced by SNAPP and normalized using edgeR-based implementation of the trimmed mean of M-values (TMM), which calculated counts per million (CPM). These normalized counts were assigned to the gene levels. Since the SHLP2 ORF was not annotated, we re-analyzed available BAM files and counted transcripts, using the summarizeOverlaps method as part of the Bioconductor package in R, that mapped back to the humanin ORF against a custom mitochondrial ORFome database. We then extracted normalized humanin transcript counts and conducted a Pearson correlation for the SHLP2 transcript against normalized counts of all genes, correcting for multiple hypothesis testing using a False Discovery Rate threshold of 0.05. We conducted a gene enrichment analysis targeted for molecular function and cellular location using ClusterProfiler Package [47].

**Targeted NAD<sup>+</sup> metabolomics.** Levels of NAD<sup>+</sup> at treatment endpoints were measured by Ultra Performance Liquid Chromatography coupled with Mass Spectrometry as previously described [48] with some modifications. Under CO<sub>2</sub> vapor, cell pellets were dissolved in chilled buffered methanol solution (75% methanol, 25% 10 mM HEPES pH 7.8) spiked with stable isotope-labeled internal standards normalization. Samples were then spun down at 16,000  $\times$  g for 10 minutes and transferred to new tubes. Pellets were saved for protein quantification as an additional normalization step. Supernatants were filtered using 4 mm 0.22  $\mu$ M syringe filters (MilliporeSigma, Burlington, MA) and lyophilized overnight with a VirTis BT6KEL-85 freeze dryer (SP Scientific Warminster, PA). Lyophilized pellets were resuspended in a 50:50 acetonitrile/10 mM ammonium bicarbonate solution with 0.1% ammonium hydroxide. 5  $\mu$ L of the extract was separated on a BEHAmide column (Waters, Milford MA) using an Acquity UPLC (Waters) and analyzed with a Xevo TQ (Waters) in multiple reaction monitoring mode (MRM). LC solvents were A: H<sub>2</sub>O with 10 mM ammonium acetate and 0.04% NH<sub>4</sub>OH. B: 95:5 Acetonitrile H<sub>2</sub>O with 10 mM ammonium acetate and 0.04% NH<sub>4</sub>OH for all metabolites. Gradient was: 5% A to 39% A at 6 minutes, to 56% A at 8 minutes, to 73% A at 8.2 minutes, and back to 5% A at 9 minutes. Each injection was 11 minutes long. Samples were run alongside an external standard curve for quantification.

#### Animals

**MPTP-lesioning and tissue collection:** C57BL/6 male mice, 8 to 10 weeks of age were purchased from Jackson Labs (Bar Harbor, Maine). 1-methyl-4-phenyl-1,2,3,6-tetrahydropyridine (MPTP) (Selleck Chemicals Co., Houston, TX) was prepared to a concentration of 5 mg/kg free-base in 0.9% saline on the day of use. C57BL/6 mice were injected with 20 mg/kg free-base MPTP. 8 days after the last injection of MPTP, mice were killed by cervical dislocation and the brain quickly removed and placed on wet ice. Striatal tissues for HPLC analysis and western immunoblotting were collected fresh corresponding to anatomical regions from bregma 1.20 to bregma 0.60, with borders dorsal to the anterior commissure, ventral to the corpus callosum, medial to the lateral ventricle, and 2.5 mm lateral from midline, and frozen at -80 °C until analysis. All procedures utilizing MPTP lesioning and tissue collection were approved by the USC IACUC.

**HPLC analysis of dopamine and its metabolites:** Dopamine concentrations were determined according to an adaptation of Irwin et al. [49] from the method of Kilpatrick et al. [50]. Tissues for analysis were homogenized in 0.4 N perchloric acid and centrifuged at 12,000  $\times$  g to pellet precipitated protein and the supernatant used for HPLC analysis. The protein pellet was resuspended in 0.5 N NaOH and the total protein concentration determined using the BCA Protein Assay (Pierce, Rockford, IL), a Biotek Model Elx800 microplate reader (Biotek Instruments, Wincoski, VT) and KCjunior software. The concentrations of dopamine were assayed by HPLC with electrochemical detection. Samples were injected with an ESA (Chelmsford, MA) autosampler. Dopamine and its metabolites were separated by a 150  $\times$  3.2 mm reverse phase 3- $\mu$ m-diameter C-18 column (ESA) regulated at 28 °C. The mobile phase MD-TM (ESA) consisted of acetonitrile in phosphate buffer and an ion-pairing agent delivered at a rate of 0.6 ml/min. The electrochemical detector was an ESA model Coularray 5600A with a four-channel analytical cell with three set potentials at -100, 50, and 220 mV. The HPLC was integrated with a Dell GX-280 computer with analytical programs including ESA Coularray for Windows software and the statistics package Prism 9 (GraphPad Software, San Diego, CA). Randomization (mixing the samples and order) of blinded samples during the process has been applied.

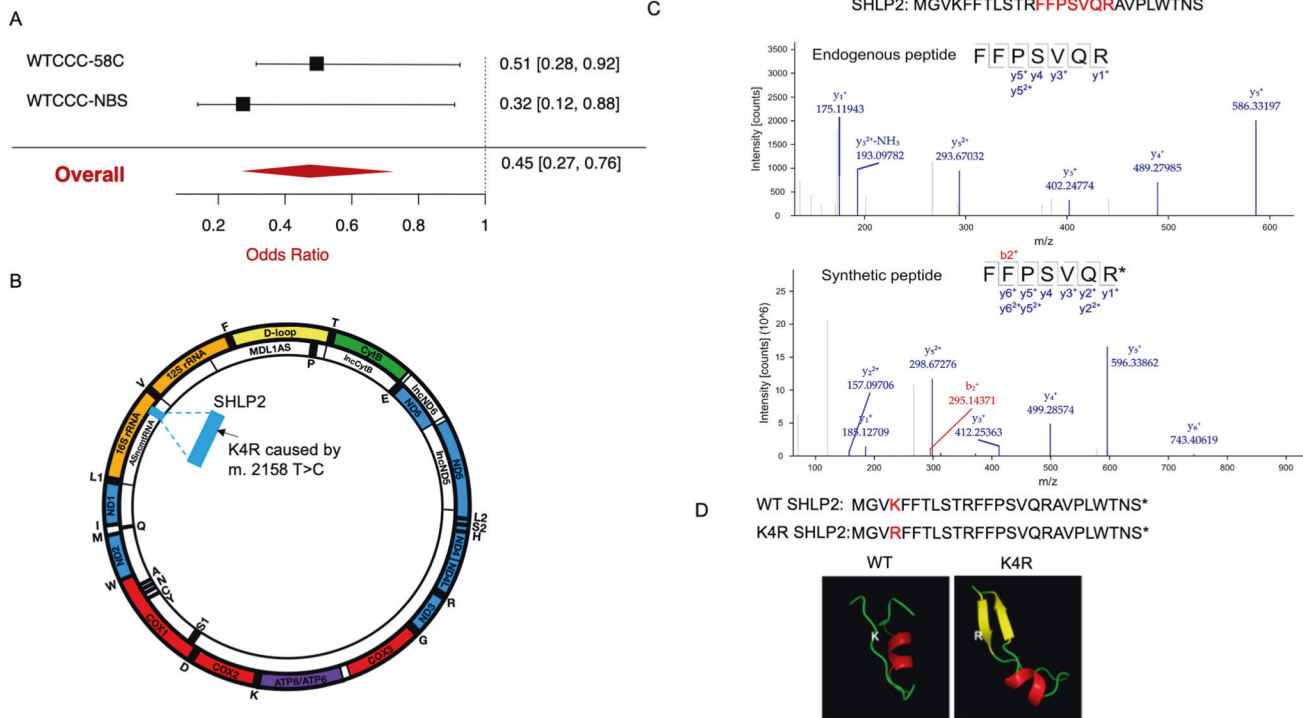
**Western Immunoblotting:** Striatal tissue samples were lysed in RIPA buffer with protease and phosphatase inhibitors (MSSAFE, Sigma), centrifuged at 16,000  $\times$  g and the soluble fraction collected for subsequent protein analysis. Total protein content was determined by BCA analysis (ThermoFisher Scientific) and 20  $\mu$ g of protein resolved on a 10% Tris-Glycine gel by electrophoresis (BioRad, Hercules, CA). Total protein was transferred to nitrocellulose membranes (BioRad), blocked (Genesee Scientific) and probed with the following antibodies overnight at 4 °C: Tyrosine hydroxylase (MAB318, Millipore Scientific, Burlington, MA) and mouse anti-beta actin (1:5000, LI-COR, Cat# 926-42212, RRID: AB\_2756372). Membranes were washed, incubated with corresponding goat anti-mouse or goat anti-rabbit conjugated near-infrared secondary fluorescent antibodies (1:5000, LI-COR, Cat# 926-32211, RRID: AB\_621843; and Cat# 926-68070, RRID: AB\_10956588), and scanned on an Odyssey imaging system (Licor, Lincoln, Nebraska, USA). Relative protein expression was quantified by optical density and normalized to beta-actin as loading control using the Licor imaging program.

**Statistical analysis:** The data is presented as the mean  $\pm$  S.E.M. Significant differences were determined by unpaired Student's *t*-tests and one-way ANOVA followed by Tukey's post hoc test. Statistical analyses were performed by using the GraphPad Prism 8 software. Values of \**p* < 0.05, \*\**p* < 0.01, \*\*\**p* < 0.001, \*\*\*\**p* < 0.0001 were considered to be statistically significant.

## RESULTS

### m.2158 T > C is associated with reduced risk of PD and with protective PD biomarkers

Mitochondria-wide association studies (MiWAS) is similar to GWAS, but instead of focusing on nuclear genetic variations, MiWAS associates mitochondrial genetic variations with particular phenotypes. Although MiWAS approaches have identified mtSNPs that associated with various diseases, the underlying mechanisms are understudied. Based on the MiWAS analysis, we pose that mitochondrial genetic variation could alter mitochondrial micro-proteins [11, 18]. Here, we tested the hypothesis that mtSNPs within small open reading frames (sORFs) would associate with PD. Previously, genetic association studies from the Welcome Trust Case Control Consortium – 1958 Birth Cohort (2,197 PD cases were compared with 2,930 controls) and the UK National Blood Service (877 PD cases were compared with 2,729 controls) identified two mtDNA variants associated with the reduced risk of PD in Caucasians. Both m.2158 T > C (Fig. 1A; an odds ratio of 0.45) and m.11251 A > G reduce the risk of PD [6]. We then mapped these two significant mtSNPs to all potential mitochondrial sORFs. Whereas m.11251 A > G is not associated with any mitochondrial sORFs, m.2158 T > C replaces amino acid Lysine (K) with amino acid Arginine (R) in the 4th position of SHLP2. Thus, we focused on the role of m.2158 T > C on SHLP2 (Fig. 1B).



**Fig. 1 Identification of SHLP2 peptide in neuronal cells and demonstration of the importance of SHLP2 SNP in PD risk.** **A** Forest plot for meta-analysis (pooled random effects) of m.2158 T > C SNP on the prevalence of PD (**B**) A schematic diagram of m.2158 T > C SNP and SHLP2 gene in a long noncoding RNA of the mitochondrial DNA. **C** Microprotein-targeted mass spectrometry identified SHLP2 in SH-SY5Y cells. (Top) MS2 scan of endogenous SHLP2 peptide from C8 enriched SH-SY5Y lysate ran in PRM mode (Bottom) MS2 scan of the spiked-in heavy labeled SHLP2 peptide ran in PRM mode. R\* represents non-radioactive heavy isotope labeling at Arginine residue. **D** Amino acid sequence of WT and K4R SHLP2 and the predicted 3D model of peptides obtained via the program PEP-FOLD3.

We examined whether m.2158 T > C in the SHLP2 sORF is associated with previously known PD risk factors. Specifically, we examined whether the m.2158 T > C is associated with plasma total cholesterol (TC), low-density lipoprotein cholesterol (LDL-C), and high-density lipoprotein cholesterol (HDL-C) because Ubiquinone-10 (coenzyme Q), a component of mitochondrial complex I, is a polyphenic molecule and is also part of the cholesterol biosynthetic pathway. Both mitochondrial complex I activity and cholesterol metabolism are significantly altered in patients with PD, and epidemiological studies showed both brain cholesterol and circulating cholesterol levels are associated with reduced PD risk and PD progression [51, 52]. To understand how the SHLP2 SNP is associated with known PD risk factors, we analyzed the association between the SHLP2 SNP and plasma cholesterol levels in control subjects (non-PD patients). We studied European ancestry individuals from three large population-based longitudinal cohorts: Health Retirement Study (HRS), Framingham Heart Study (FHS), and Cardiovascular Health Study (CHS). Based on the analysis, the alternative allele, m.2158 C is associated with higher TC (HRS: beta = 8.9,  $p = 0.025$ ; FHS: beta = 12.77,  $p = 0.01$ ) and LDL-C (HRS: beta = 3.71,  $p = 0.43$ ; FHS: beta = 9.37,  $p = 0.01$ ) levels while no association was found with HDL-C (HRS: beta = 1.00,  $p = 0.48$ ; FHS: beta = 0.37,  $p = 0.85$ ) levels (Table 1, Supplementary Tables 1–4).

### Endogenous SHLP2 is detectable by mass spectrometry in a neuronal cell line

Endogenous SHLP2 was detected by the microprotein targeted proteomics approach, which enriched microproteins using a solid-phase extraction with C8 cartridges, followed by digestion with trypsin, and liquid chromatography-mass spectrometry [43]. First, we used synthetic light and heavy isotope-labeled SHLP2 peptides to develop a targeted parallel reaction monitoring assay. This

**Table 1.** Meta-analysis on m. 2158 T > C polymorphism and lipids.

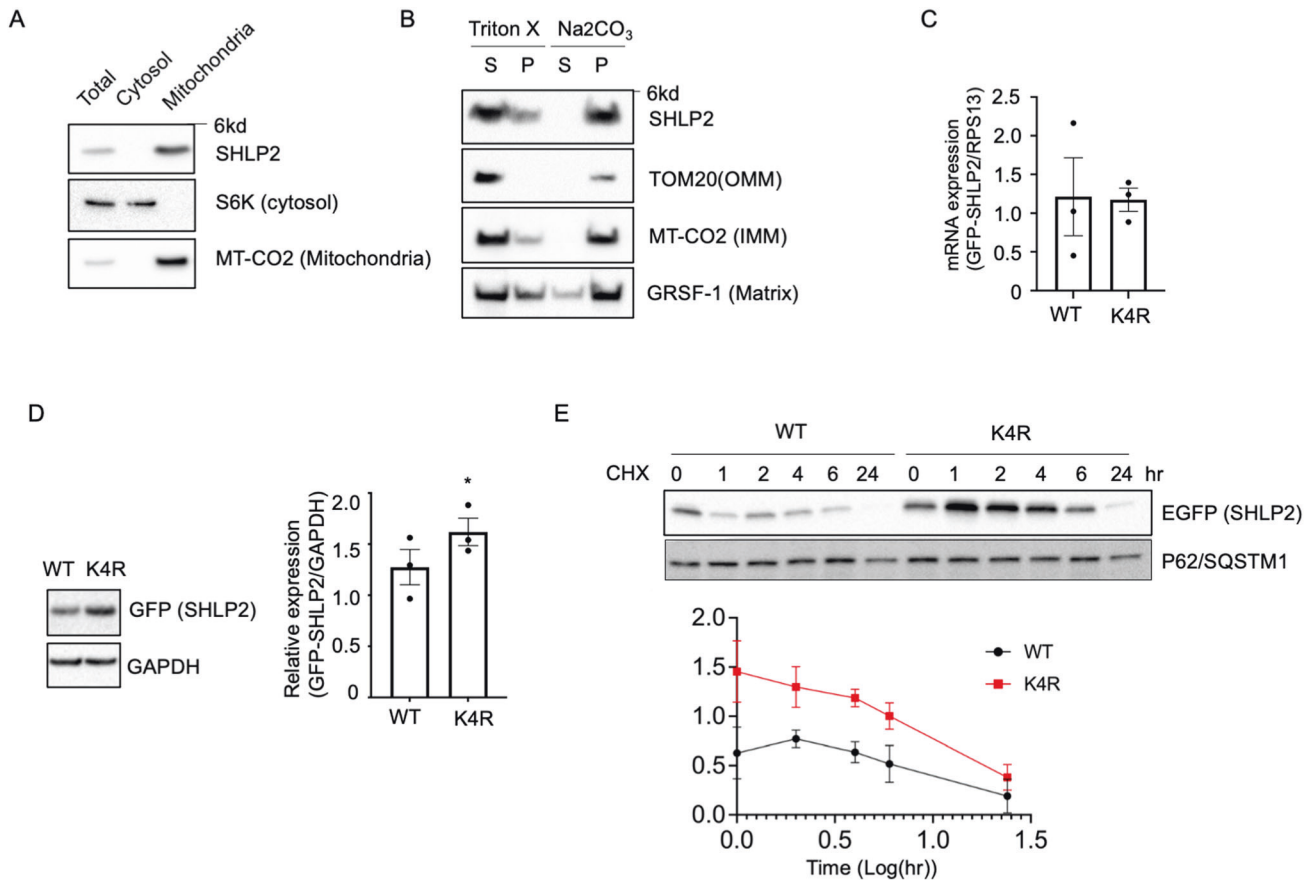
	Effect	StdErr	P value	Direction
TC	8.14	2.85	0.004	++–
LDL	5.76	2.7	0.033	+++
HDL	0.39	1.06	0.715	++–

analysis identified a detectable tryptic peptide, FFPSVQRA, from SHLP2. Next, we analyzed lysates from SH-SY5Y cells with or without spiked the heavy isotope-labeled SHLP2 peptide and successfully identified endogenous SHLP2 expression (Fig. 1C). This data provides experimental evidence for SHLP2 translation from the non-coding RNA region of the mitochondrial DNA.

### Structural prediction of the K4R SHLP2

The analytical package PROVEAN (PROtein Variation Effect Analyzer; <http://provean.jcvi.org>) [53, 54] was used to predict the potential impact of the K4R substitution in SHLP2. The PROVEAN score for SHLP2 K4R replacement was -3.0, which is below the predicted cutoff score ( $= -2.5$ ). Because the PROVEAN score is smaller than the given threshold, K4R SHLP2 is predicted to produce a functional change. Thus, we compared to the predicted structures of SHLP2 with K4R variant using the de novo modeling algorithm PEP-FOLD3 [55]. The predicted 3D structure of WT SHLP2 showed a loop-like structure with a small alpha-helix in the middle region, suggesting a flexible conformation. The K4R substitution potentially results in a significant alteration in the predicted 3D structure, forming an additional beta-sheet structure (Fig. 1D). Additionally, the change in the number of potential hydrogen bond from 3 in arginine to 2 in lysine could account for altered structure and function of the K4R SHLP2.





**Fig. 2 SHLP2 localization and stability.** **A** Western blot of endogenous SHLP2 from subcellular fractions of HEK293 cells. **B** Extraction of endogenous SHLP2 from HEK293 mitochondria with buffer containing 1% Triton X-100 or Na<sub>2</sub>CO<sub>3</sub>. **C–E** HEK293 cells were transiently transfected with WT and K4R SHLP2 – tagged with EGFP. **C** mRNA expression of WT and K4R SHLP2 measured by qRT-PCR (**D**) (left) Representative image of total protein expression of WT and K4R SHLP2 (right) Quantification of the expression (**E**) Pulse-chase experiment of WT and K4R SHLP2. Cells were treated with 100ug/ml cycloheximide for the indicated time. p62/SQSTM1 is used as a loading control. Data are reported as mean ± SEM (n = 3/per group). Significant differences were determined by Student's t-tests. \*p < 0.05.

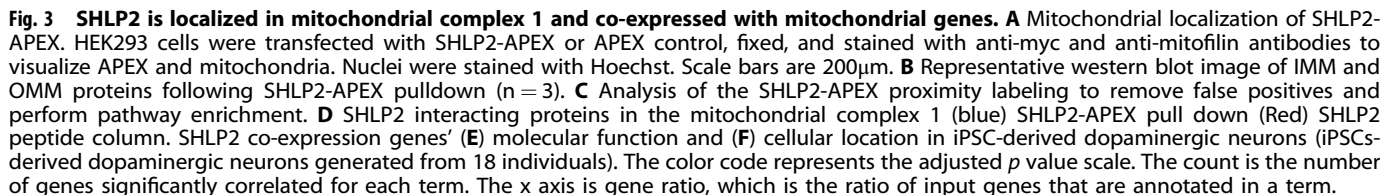
### Subcellular localization and stability of SHLP2 and K4R variant

As protein localization in a cell is strongly associated with its function, we analyzed the putative subcellular localization of SHLP2 by using TargetP - 2.0 [56] and DeepLoc [57], a protein subcellular localization prediction algorithm based on deep learning. SHLP2 was predicted to be a mitochondrial membrane integrated protein. Indeed, with a custom antibody against full length SHLP2, SHLP2 is enriched in the mitochondria fraction, but not the cytoplasmic fraction (Fig. 2A). Sodium carbonate extraction is regarded as a canonical way to distinguish integral membrane proteins from other membrane-associated proteins and soluble proteins. SHLP2 and other mitochondrial membrane proteins (TOM20 and MT-CO2) were resistant to sodium carbonate extraction suggesting SHLP2 is an integral membrane protein (Fig. 2B). The SHLP2 SNP causes an amino acid substitution of lysine with arginine in SHLP2. Lysine is a hotspot of post-translational modification, including ubiquitination, acetylation, and sumoylation. These post-translational modifications play essential roles in protein subcellular localization and/or stability. We thus examined the impact of the amino acid substitution on protein stability. Creating an isogenic cell line using mtDNA editing could be an ideal approach to answer this question. Despite recent advances in mtDNA editing, the editing frequencies are still limited, up to 50%, and there is still off-target editing [58]. We thus overexpressed cells with plasmids carrying WT SHLP2 and K4R variant tagged with EGFP at the C-terminus and examined the stability of the peptides. HEK293 cells were

transiently transfected with plasmids carrying either WT SHLP2 or K4R variant tagged with EGFP. Even though the mRNA expression of both WT and K4R SHLP2 was similar in HEK293 cells (Fig. 2C), the K4R SHLP2 showed higher fully expressed protein levels than WT (Fig. 2D). Furthermore, cycloheximide-treated pulse-chase experiments showed K4R SHLP2 is more stable than WT SHLP2 (Fig. 2E).

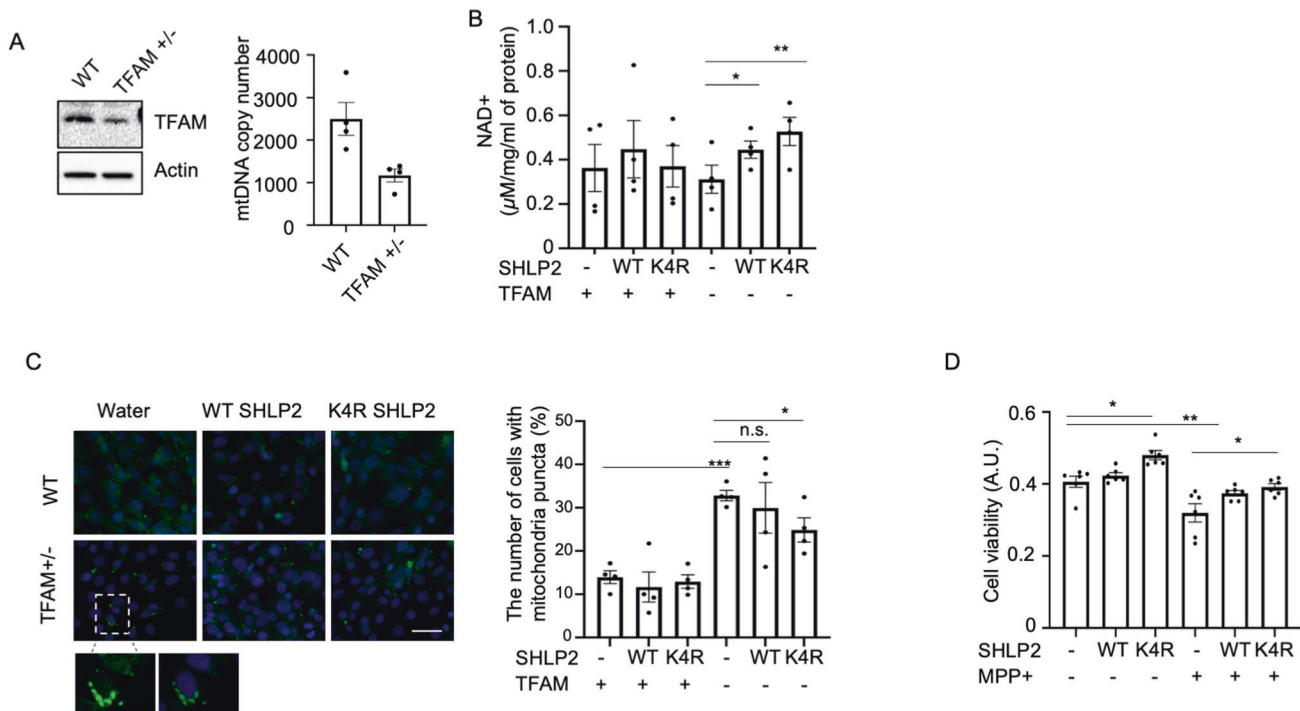
### SHLP2 localizes to mitochondrial complex 1

To understand the molecular function of SHLP2, we sought to identify SHLP2 interacting proteins. The identification of microprotein-protein interactions has proven to be a successful approach for the functional characterization of novel microproteins. However, traditional immunoprecipitation methods result in the enrichment of nonspecific interactions for microproteins. On the other hand, an in situ proximity tagging method that relies on an engineered ascorbate peroxidase 2 (APEX2) has been a useful tool to identify microprotein-protein interactions [45]. In this approach, APEX is fused to a protein of interest. Expression of the APEX fusion protein followed by treatment of cells with hydrogen peroxide in the presence of biotin-phenol covalently labels proteins proximal to the APEX fusion proteins with biotin. These biotinylated proteins can then be enriched and analyzed by mass spectrometry and western blotting, and because proteins adjacent to the APEX fusion protein are preferentially biotinylated, the resulting mass spectrometry data provide a readout of the protein environment around the fusion protein. Thus, we first utilized



Application of SAINT to the SHLP2-APEX data set results in 597 proteins (Fig. 3C) with an average spectral count of greater than 5. Then, the enrichment analysis of proteins labeled by SHLP2-APEX using String-wikiPathways (v. 11.5) found proteins are enriched in mitochondrial complex 1 and TCA cycle (Fig. 3C). A previous mitochondrial matrix-targeted APEX study showed that mitochondrial matrix proteome from APEX labeling contain both soluble matrix proteins and inner mitochondrial membrane (IMM) proteins that contact the matrix space [60]. Likewise, mitochondrial membrane protein SHLP2-APEX appear to label both IMM proteins as well as mitochondrial matrix protein. In addition to APEX proximity labeling, we created a SHLP2 peptide column and a scrambled peptide column as a control to identify direct SHLP2 binding proteins (Supplemental Fig. 1A). This antibody-free approach helped to reduce the background interactions from antibody non-specificity while pulling down direct interacting partners of SHLP2. We added isolated mitochondrial protein extract and pulled down potential binding proteins. The elution was analyzed by mass spectrometry. We identified an enrichment of SHLP2 binding partner in mitochondrial complex 1 and TCA cycle proteins (Supplemental Fig. 1B). After analyzing data from





**Fig. 4** K4R SHLP2, produced by individuals who carry the m. 2158 T > C polymorphism, has superior protection against mitochondrial dysfunction. **A** TFAM expression and mtDNA copy number **(B)** NAD<sup>+</sup> level was measured by mass spectrometry in WT and TFAM heterozygous knockout MEFs treated with 2.5 μM WT and K4R SHLP2 for 24 hr. **C** (left) Immunostaining of mitochondria using COX IV antibodies. (right) The number of cells carrying mitochondria puncta was quantified. A total of 500–700 cells per group from four representative images were quantified. Scale bar is 50 μm. **D** Cell viability was measured by MTT assay in SH-SY5Y cells treated with 2.5 μM WT or K4R SHLP2 in the presence of 200 μM MPP<sup>+</sup>. Data are reported as mean ± SEM (n = 4/per group). Significant differences were determined by Student's t-tests. \**p* < 0.05, \*\**p* < 0.01, \*\*\**p* < 0.001, n.s. = non significant.

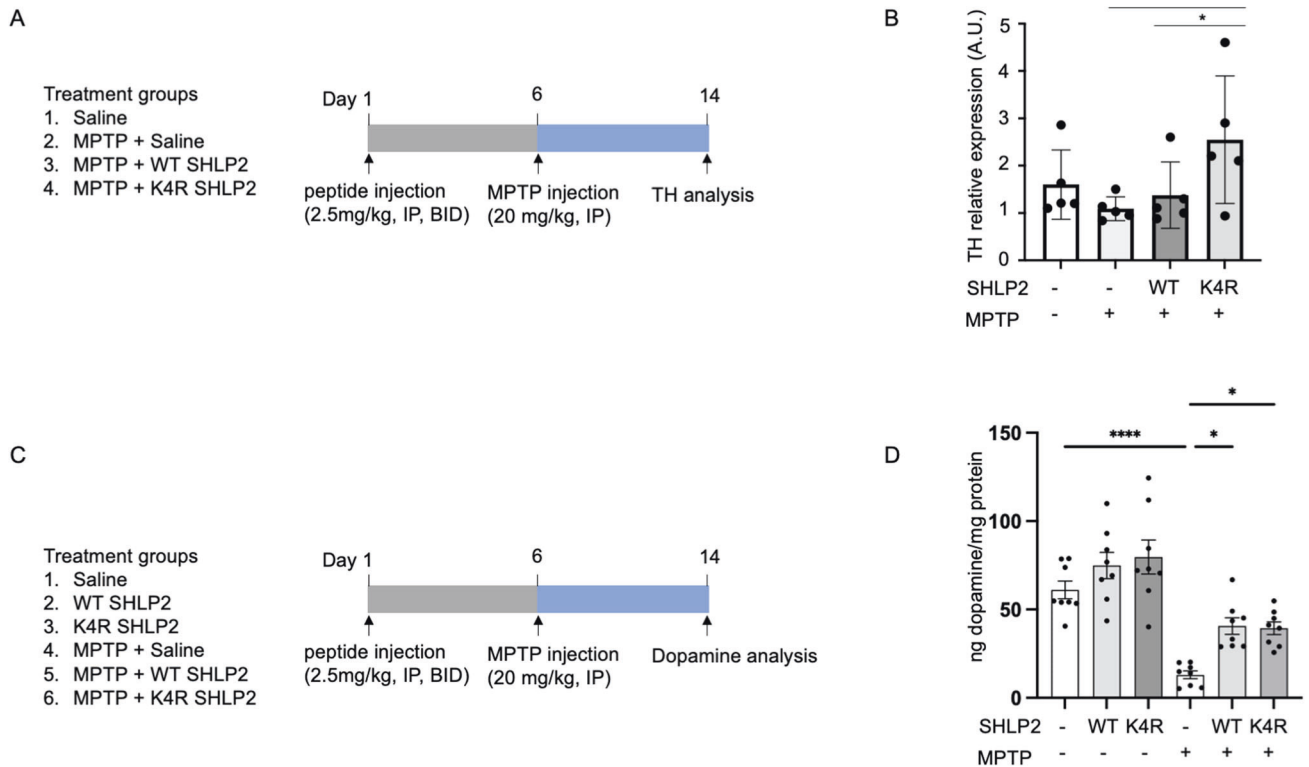
APEX proximity labeling and the SHLP2 peptide column, we concluded that SHLP2 directly binds to the Pd modules of mitochondrial complex 1 (Fig. 3D). Mitochondrial complex 1 consists of four modules - N module (NADH oxidation), Q module (reduction of ubiquinone), Pd and Pp modules (proton pumping). These results suggest SHLP2 might play important roles in mitochondrial complex 1 related activity such as proton pumping.

To understand the molecular function of SHLP2 in neurons, we carried out a co-expression analysis on publicly available RNA-seq data derived from iPSCs-derived dopaminergic neurons [61]. Interestingly, significant genes that co-expressed with SHLP2 were related to NADH dehydrogenase activity (Fig. 3E) and the inner mitochondrial membrane (Fig. 3F); these associations align with our experimental data showing SHLP2 interaction with mitochondrial complex 1.

#### SHLP2 protects against mitochondrial dysfunction in vitro

Because we determined SHLP2 is in mitochondrial complex 1, we examined whether SHLP2 modifies mitochondrial function and prevents mitochondrial dysfunction. TFAM is a mitochondrial DNA binding protein essential for mitochondrial gene expression and genome maintenance. Patients with PD show reduced levels of TFAM and mtDNA copy number [62, 63]. Using TFAM heterozygous knockout mouse embryonic fibroblasts (MEFs), the potential protective role of SHLP2 against mitochondrial dysfunction was examined. Before we treated cells with synthetic SHLP2 peptide, we examined whether synthetic SHLP2 peptide enters to the mitochondria. First, MitoFates [64] predicted not only mitochondrial localization of SHLP2 but also a TOM20 recognition motif (amino acids 16–20) of SHLP2. Indeed, a SHLP2 column experiment for identification of SHLP2 binding proteins revealed that synthetic SHLP2 binds to TOM20 as well as other proteins in

the mitochondrial protein import system. The results showed SHLP2 potentially binds to TOMM34, a chaperone to recruit proteins to transporters in outer mitochondrial membrane; SHLP2 binds to TOM20, then intermembrane space proteins (Timm9, Timm10, Timm10B), then is finally recruited to Tim50 in the inner mitochondrial membrane (Supplementary Fig. 2A). Moreover, rapid mitochondria immunoprecipitation showed SHLP2 was localized in the mitochondria after 30 min of treatment (Supplementary Fig. 2B). Based on these three approaches, we concluded that synthetic SHLP2 can be localized into the mitochondria and tested the hypothesis that synthetic SHLP2 peptide treatment protects against mitochondrial dysfunction found in TFAM heterozygous knockout MEFs. TFAM heterozygous knockout MEFs showed a 50% reduction in mtDNA copy number as reported elsewhere (Fig. 4A). Administration of WT SHLP2 and K4R SHLP2 did not alter levels of TFAM and Peroxisome proliferator-activated receptor-gamma coactivator (PGC-1 α) protein, the master regulator of mitochondrial biogenesis (Supplemental Fig. 3A–3C). Because mitochondrial complex 1 is involved in NADH oxidation, we examined the levels of NAD<sup>+</sup> in TFAM heterozygous knockout MEFs. Exposure of MEFs cells to K4R SHLP2 resulted in increased NAD<sup>+</sup> levels (Fig. 4B). Mitochondrial lysosome crosstalk has been recognized in previous studies via NAD<sup>+</sup>. TFAM knockout reduces NAD<sup>+</sup>, thereby including lysosomal dysfunction and aberrant mitochondrial morphology and increases mitochondrial mass due to the defective autophagosome-lysosome fusion [65]. Because K4R SHLP2 increased NAD<sup>+</sup> levels in TFAM knockout cells, we examined the effect of K4R SHLP2 on mitochondrial morphology in the same cells. K4R SHLP2 significantly reduced the number of cells carrying enlarged mitochondria (Fig. 4C). Taken together, K4R SHLP2 can mitigate the mitochondrial dysfunction during TFAM reduction.



**Fig. 5** Pretreatment of mice with K4R SHLP2 shows protection in MPTP-lesioned mice. **A** Schematic diagram of treatment. **B** Tyrosine hydroxylase (TH) expression in the striatum from mice treated with or without MPTP and WT SHLP2 and K4R variant. (n = 5 per group). **C** Schematic diagram of treatment. **D** Dopamine contents were measured by using high-performance liquid chromatography with electrochemical detection in the striatum from mice treated with or without MPTP and WT and K4R SHLP2 (n = 8 per group). Data are reported as mean  $\pm$  SEM. Significant differences were determined by one-way ANOVA followed by Tukey's post hoc test  $^*p < 0.05$   $^{**} < 0.01$   $^{****} < 0.0001$ .

**SHLP2 has protection in SH-SY5Y cells treated with MPP +**  
Reduced mitochondrial complex I activity has been found in PD pathology [66]. Mitochondrial complex I inhibitors such as MPTP and its metabolite MPP+ lead to neuronal cell death in vitro and in vivo, which is widely used for cellular and animal PD models [3, 67, 68]. Using SH-SY5Y cells treated with MPP +, the protective effect of WT SHLP2 and K4R variant in MPP + -induced cell death in vitro was examined. K4R SHLP2 treatment reduced the MPP + -induced cell death in SH-SY5Y cells (Fig. 4D).

#### SHLP2 has protection in an MPTP mice model of PD

To examine the potential role of WT SHLP2 and K4R variant to provide protection against neurotoxin-induced dopamine-depletion in vivo, C57BL/6 mice exposed to MPTP were also administered WT SHLP2 or K4R variant starting 5 days prior to MPTP-induced cell death (Fig. 5A). Pretreatment of SHLP2 peptides could recapitulate the SHLP2 production in humans, which individuals carrying m.2158 T allele or m.2158 C allele produce WT SHLP2 or K4R variant, respectively. MPTP administration results in the degeneration of dopaminergic neurons within the substantia nigra, a brain region known for its projection to the striatum, leading to a consequential reduction in striatal dopamine levels. The assessment of dopaminergic neuron viability and functionality hinges on the quantification of tyrosine hydroxylase (TH) expression, a pivotal enzyme in dopamine biosynthesis [4]. To investigate the potential protective effects of SHLP2 against the loss of dopaminergic neurons triggered by MPTP, we performed western blot analysis to determine the TH expression levels in the striatum. Administration of K4R SHLP2 (2.5 mg/kg, IP, BID) protected against the loss of striatal TH expression in MPTP-treated mice while WT SHLP2 did not provide a similar degree of

protection (Fig. 5B). Then, the dopamine levels in the striatum were measured by employing reverse-phase HPLC-electrochemical detection to determine whether SHLP2 can prevent the MPTP-induced reductions in dopamine (Fig. 5C). MPTP (20 mg/kg) induces a significant decrease in dopamine levels in the striatum of injected mice (Fig. 5D). WT and K4R SHLP2 administration rescues the MPTP-induced dopamine loss in the striatum compared to MPTP treated mice (Fig. 5D).

#### DISCUSSION

We determined that the m.2158 T > C polymorphism, which reduces PD risk, alters the MDP SHLP2 by mutating its fourth amino acid (K4R). We detected endogenous SHLP2 in neuronal cells using mass spectrometry. Through experimentation, we showed that SHLP2 microprotein was protective against mitochondrial dysfunction in vitro and in a mitochondrial toxin model of PD in mice.

Studies on the functional role of mtDNA polymorphisms have mainly focused on the 13 mtDNA-encoded respiratory chain proteins and free radical production [69, 70]. However, emerging studies have showed that small open reading frames within mtDNA also produces microproteins [9]. Here, we showed that PD protective SNP, m.2158 T > C affects a mitochondrial microprotein encoded from a long noncoding RNA of the mtDNA. m.2158 T > C changes the amino acid of the microprotein SHLP2 from lysine to arginine (K4R). We further demonstrated a protective role of K4R SHLP2 in in vitro and in vivo models of PD. We thus provided insights on the significance of the study investigating the functional role of mtSNPs on MDPs, which potentially determine the potential therapeutic targets. In addition, our work

demonstrates how interpretation of previous associative studies with mtSNPs can be re-interpreted with a microprotein focus.

We identified SHLP2 by mass spectrometry in neuronal cells, which supports the coding potential of long noncoding RNA of the mtDNA. Traditional mass spectrometry often loses small proteins during sample preparation. Recent advancements in microprotein enrichment in mass spectrometry enabled us to identify the mitochondrial microprotein, SHLP2 in SH-SY5Y neuronal cells. Because peptidase causes microproteins to degrade rapidly, we began our sample preparation with a large quantity of starting materials. Next, we needed to detect microproteins by applying an enrichment. In previous research, four enrichment techniques have been proposed: acidic precipitation, 30 kDa MWCO filter, polyacrylamide gel electrophoresis, and reverse-phase (C8) column enrichment [43]. C8 column enrichment was originally developed to enrich plasma and tissue extracts for peptide hormones by removing larger molecular weight proteins before measurement by radioimmunoassay. Applying this method to enrich the lower molecular weight proteins gave excellent results to detect SHLP2 in SH-SY5Y neuronal cells.

We found SHLP2 is localized in mitochondrial complex 1, particularly the proton pumping module. K4R SHLP2 could bind to mitochondrial complex 1 more stably and prevent the decline of mitochondrial complex 1 activity (e.g., decline of NAD<sup>+</sup> levels). Indeed, K4R SHLP2 treatment increased NAD<sup>+</sup> levels in the TFAM heterozygous knockout MEFs. Given that declines in NAD<sup>+</sup> have been shown to be harmful, we propose that an increased level of NAD<sup>+</sup> by K4R SHLP2 could serve as a protective mechanism of action. A decline in NAD<sup>+</sup> levels has been shown in PD patients, in vitro PD models, and in vivo PD models [71–74]. The blood of idiopathic PD patients revealed a decrease in the NAD<sup>+</sup>/NADH ratio [75]. The iPSC-derived neurons that express LRRK2 G2019S or mutant GBA1 showed mitochondrial dysfunction and a significant reduction of NAD<sup>+</sup> levels [74]. Finally, a PD mouse model that was related to MPTP also showed reduction of NAD<sup>+</sup> in the striatum and ventral midbrain but not in the frontal cortex [72]. In contrast, strategies to boost NAD<sup>+</sup> metabolism, including the use of NAD<sup>+</sup> precursors (e.g., NMN and NR) or small molecules that increase the rate of NAD<sup>+</sup> biosynthesis or inhibit NAD<sup>+</sup> degradation, have been used in current clinical trials [71]. Like NAD<sup>+</sup> targeting therapies, we showed K4R SHLP2 not only increases NAD<sup>+</sup> levels but also protected against MPTP mice model of PD.

A clinical study showed NAD<sup>+</sup> precursors increase not only NAD<sup>+</sup> levels but also total cholesterol and LDL [76]. Our meta-analysis showed that the PD-protective SHLP2 SNP is also associated with higher total cholesterol and LDL. Further studies examining NAD<sup>+</sup> and its related metabolites levels in individuals carrying the SHLP2 SNP will confirm our observation in cells that SHLP2 increases NAD<sup>+</sup> levels. These future studies could provide an insight on pharmacodynamics markers in development of PD therapeutics in the form of SHLP2 analogs. NAD<sup>+</sup> was reduced in PD patients derived iPSCs carrying common mutations found in PD (LRRK2, GBA1) [73, 74]. The genetic interaction between SHLP2 SNP and PD mutations in the risk of PD and protective effects of K4R SHLP2 in PD patients-derived iPSCs carrying the PD mutations could be an essential direction to explore further in the future. Studies showed specific genetic mutations can manifest different responses to pharmacological therapies [77]. The future studies on the interaction between SHLP2 variant and PD mutations could be informative to develop SHLP2's analog for personalized medicine.

In summary, we provide evidence that mitochondria microproteins play a central role in understanding disease-associated mtSNPs. We have taken a published mtSNP association that lacked a mechanism and have been able to demonstrate one possible reason for this association. We found that the mitochondrial inner membrane microprotein SHLP2 is affected by the m.2158 T>C polymorphism associated with reduced risk of PD. K4R SHLP2

produced by individuals carrying the SNP have higher cellular stability compared to WT SHLP2. Our data suggests that SHLP2 and its variants are PD protective microproteins and that administration of K4R may be able to protect against PD. Therefore, a deeper understanding of the effects of this genetic polymorphism and its related mitochondrial microprotein will provide a basis for developing novel therapeutic targets in PD.

## DATA AVAILABILITY

The SHLP2 mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via MassIVE (<http://proteomecentral.proteomexchange.org>) via the PRIDE partner repository with the data set identifier PXD034482. The APEX and SHLP2 column mass spectrometric raw data are deposited at <ftp://MSV0000089658@massive.ucsd.edu> (password: Shlp2\_2022) with the MassIVE ID MSV0000087928. (after publication: <ftp://massive.ucsd.edu/MSV0000089658/>).

## REFERENCES

- Gao X-Y, Yang T, Gu Y, Sun X-H. Mitochondrial Dysfunction in Parkinson's Disease: From Mechanistic Insights to Therapy. *Front aging Neurosci.* 2022;14:885500–885500.
- Gonzalez-Rodriguez P, Zampese E, Stout KA, Guzman JN, Ilijic E, Yang B, et al. Disruption of mitochondrial complex I induces progressive parkinsonism. *Nature.* 2021;599:650–6.
- Onofri M, Ghilardi MF. MPTP induced Parkinsonian syndrome: long term follow-up and neurophysiological study. *Ital J Neurological Sci.* 1990;11:443–58.
- Mustapha M, Taib CNM. MPTP-induced mouse model of Parkinson's disease: A promising direction for therapeutic strategies. *Bosn J basic Med Sci.* 2021;21:422–33.
- van der Walt JM, Nicodemus KK, Martin ER, Scott WK, Nance MA, Watts RL, et al. Mitochondrial polymorphisms significantly reduce the risk of Parkinson disease. *Am J Hum Genet.* 2003;72:804–11.
- Hudson G, Nalls M, Evans JR, Breen DP, Winder-Rhodes S, Morrison KE, et al. Two-stage association study and meta-analysis of mitochondrial DNA variants in Parkinson disease. *Neurology.* 2013;80:2042–8.
- Ross OA, McCormack R, Maxwell LD, Duguid RA, Quinn DJ, Barnett YA, et al. mt4216C variant in linkage with the mtDNA T1 cluster may confer a susceptibility to mitochondrial dysfunction resulting in an increased risk of Parkinson's disease in the Irish. *Exp Gerontol.* 2003;38:397–405.
- Gaweda-Walerych K, Maruszak A, Safranow K, Bialecka M, Kłodowska-Duda G, Czyżewski K, et al. Mitochondrial DNA haplogroups and subhaplogroups are associated with Parkinson's disease risk in a Polish PD cohort. *J Neural Transm.* 2008;115:1521–6.
- Miller B, Kim S-J, Kumagai H, Yen K, Cohen P. Mitochondria-derived peptides in aging and healthspan. *J Clin Investig.* 2022;132:1–9.
- Kim SJ, Miller B, Kumagai H, Silverstein AR, Flores M, Yen K. Mitochondrial-derived peptides in aging and age-related diseases. *Geroscience* 2020.
- Miller B, Kim S-J, Mehta HH, Cao K, Kumagai H, Thumay N, et al. Mitochondrial DNA variation in Alzheimer's disease reveals a unique microprotein called SHMOOSE. *Molecular psychiatry* 2022.
- Cobb LJ, Lee C, Xiao J, Yen K, Wong RG, Nakamura HK, et al. Naturally occurring mitochondrial-derived peptides are age-dependent regulators of apoptosis, insulin sensitivity, and inflammatory markers. *Aging (Albany NY).* 2016;8:796–809.
- Nashine S, Cohen P, Nesburn AB, Kuppermann BD, Kenney MC. Characterizing the protective effects of SHLP2, a mitochondrial-derived peptide, in macular degeneration. *Sci Rep.* 2018;8:15175.
- Okada AK, Teranishi K, Lobo F, Isas JM, Xiao J, Yen K, et al. The Mitochondrial-Derived Peptides, HumaninS14G and Small Humanin-like Peptide 2, Exhibit Chaperone-like Activity. *Sci Rep.* 2017;7:7802.
- Xiao J, Howard L, Wan J, Wiggins E, Vidal A, Cohen P, et al. Low circulating levels of the mitochondrial-peptide hormone SHLP2: novel biomarker for prostate cancer risk. *Oncotarget.* 2017;8:94900–9.
- Ivana R et al. Plasma mitochondrial derived peptides MOTs-c and SHLP2 positively associate with android and liver fat in people without diabetes. *Biochimica et Biophysica Acta (BBA) - General Subjects* 2021;1865.
- Yen K, Wan J, Mehta HH, Miller B, Christensen A, Levine ME, et al. Humanin Prevents Age-Related Cognitive Decline in Mice and is Associated with Improved Cognitive Age in Humans. *Sci Rep.* 2018;8:14212–10.
- Zempo H, Kim S-J, Fuku N, Nishida Y, Higaki Y, Wan J et al. A pro-diabetogenic mtDNA polymorphism in the mitochondrial-derived peptide, MOTs-c. *Aging (Albany, NY)* 2021;13.
- HRS. HRS, Sample Sizes and Response Rates. *University of Michigan* 2011: 1–13.



20. Juster FT, Suzman R. An Overview of the Health and Retirement Study. *J Hum Resour.* 1995;30:57–556.
21. dbGaP. dbGaP, Health and Retirement Study. *National Center for Biotechnology Information: Bethesda, MD* 2012.
22. Laurie CC, Doherty KF, Mirel DB, Pugh EW, Bierut LJ, Bhargava T, et al. Quality control and quality assurance in genotypic data for genome-wide association studies. *Genet Epidemiol.* 2010;34:591–602.
23. HRS. HRS, Quality control report for genotypic data. *University of Washington: St Louis, MO* 2012:1–44.
24. Chang CC, Chow CC, Tellier LCAM, Vattikuti S, Purcell SM, Lee JJ. Second-generation PLINK: Rising to the challenge of larger and richer datasets. *Giga-science.* 2015;4:7–7.
25. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, et al. PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. *Am J Hum Genet.* 2007;81:559–75.
26. Crimmins E Documentation of biomarkers in the 2006 and 2008 Health and Retirement Study. *Ann Arbor, MI: Survey Research Center University of Michigan* 2013.
27. Crimmins E, Kim JK, McCreath H, Faul J, Weir D, Seeman T. Validation of Blood-Based Assays Using Dried Blood Spots for Use in Large Population Studies. *Bio-demography Soc Biol.* 2014;60:38–48.
28. Duncan MS, Vasan RS, Xanthakis V. Trajectories of Blood Lipid Concentrations Over the Adult Life Course and Risk of Cardiovascular Disease and All-Cause Mortality: Observations From the Framingham Study Over 35 Years. *J Am Heart Assoc.* 2019;8:e011433–e011433.
29. EM Crimmins, J Faul, B Thyagarajan, DR Weir. Health and Retirement Study Website. University of Michigan. [http://hrsonline.isr.umich.edu/modules/meta/vbs/2016/desc/HRS2016VBSDD.pdf?\\_ga=2.171951269.408614381.1589823285-975920407.1589823285](http://hrsonline.isr.umich.edu/modules/meta/vbs/2016/desc/HRS2016VBSDD.pdf?_ga=2.171951269.408614381.1589823285-975920407.1589823285) Venous blood collection and assay protocol in the 2016 Health and Retirement Study Venous Blood Study (VBS). 2017.
30. Tian C, Gregersen PK, Seldin MF. Accounting for ancestry: population substructure and genome-wide association studies. *Hum Mol Genet.* 2008;17:R143–R150.
31. Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D. Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet.* 2006;38:904–9.
32. Patterson N, Price AL, Reich D. Population structure and eigenanalysis. *PLoS Genet.* 2006;2:2074–93.
33. Fried LP, Borhani NO, Enright P, Furberg CD, Gardin JM, Kronmal RA, et al. The Cardiovascular Health Study: design and rationale. *Ann Epidemiol.* 1991;1:263–76.
34. Dawber TR, Meadors GF, Moore FE Jr. Epidemiological approaches to heart disease: the Framingham Study. *Am J Public Health Nations Health.* 1951;41:279–81.
35. Feinleib M, Kannel WB, Garrison RJ, McNamara PM, Castelli WP. The Framingham Offspring Study. Design and preliminary data. *Prev Med.* 1975;4:518–25.
36. Splansky GL, Corey D, Yang Q, Atwood LD, Cupples LA, Benjamin EJ, et al. The Third Generation Cohort of the National Heart, Lung, and Blood Institute's Framingham Heart Study: design, recruitment, and initial examination. *Am J Epidemiol.* 2007;165:1328–35.
37. Willer CJ, Schmidt EM, Sengupta S, Peloso GM, Gustafsson S, Kanoni S, et al. Discovery and refinement of loci associated with lipid levels. *Nat Genet.* 2013;45:1274–83.
38. Conomos MP, Miller MB, Thornton TA. Robust Inference of Population Structure for Ancestry Prediction and Correction of Stratification in the Presence of Relatedness. *Genet Epidemiol.* 2015;39:276–93.
39. Ding J, Sidore C, Butler TJ, Wing MK, Qian Y, Meirelles O, et al. Assessing Mitochondrial DNA Variation and Copy Number in Lymphocytes of ~2,000 Sardinians Using Tailored Sequencing Analysis Tools. *PLoS Genet.* 2015;11:e1005306.
40. Liu C, Fetterman JL, Qian Y, Levy D, NHLBL Trans-Omics for Precision Medicine TOPMed Consortium. Presence and transmission of mitochondrial heteroplasmic mutations in human populations of European and African ancestry. *Mitochondrion.* 2021;60:33–42.
41. Liu C, Fetterman JL, Liu P, Luo Y, Larson MG, Vasan RS, et al. Deep sequencing of the mitochondrial genome reveals common heteroplasmic sites in NADH dehydrogenase genes. *Hum Genet.* 2018;137:203–13.
42. Shipley MM, Mangold CA, Szpara ML. Differentiation of the SH-SY5Y human neuroblastoma cell line. *J visualized Exp.* 2016;2016:53193–53193.
43. Ma J, Diedrich JK, Jungreis I, Donaldson C, Vaughan J, Kellis M, et al. Improved Identification and Analysis of Small Open Reading Frame Encoded Polypeptides. *Anal Chem (Wash).* 2016;88:3967–75.
44. Chen WW, Freinkman E, Sabatini DM. Rapid immunopurification of mitochondria for metabolite profiling and absolute quantification of matrix metabolites. *Nat Protoc.* 2017;12:2215–31.
45. Chu Q, Rathore A, Diedrich JK, Donaldson CJ, Yates JR, Saghatelian A. Identification of Microprotein-Protein Interactions via APEX Tagging. *Biochem (Easton).* 2017;56:3299–306.
46. Tan B, Peng S, Yatim SMJM, Gunaratne J, Hunziker W, Ludwig A. An Optimized Protocol for Proximity Biotinylation in Confluent Epithelial Cell Cultures Using the Peroxidase APEX2. *STAR Protoc.* 2020;1:100074–100074.
47. Yu G, Wang L-G, Han Y, He Q-Y. clusterProfiler: an R Package for Comparing Biological Themes Among Gene Clusters. *Omics (Larchmont, NY).* 2012;16:284–7.
48. Trammell SAJ, Brenner C. Targeted, LCMS-based metabolomics for quantitative measurement of NAD<sup>+</sup> metabolites. *Comput Struct Biotechnol J.* 2013;4:e201301012–e201301012.
49. Irwin I, Finnegan KT, Delaney LE, Di Monte D, Langston JW. The relationships between aging, monoamine oxidase, striatal dopamine and the effects of MPTP in C57BL/6 mice: a critical reassessment. *Brain Res.* 1992;572:224–31.
50. Kilpatrick IC, Jones MW, Phillipson OT. A Semiautomated Analysis Method for Catecholamines, Indoleamines, and Some Prominent Metabolites in Micro-dissected Regions of the Nervous System: An Isocratic HPLC Technique Employing Coulometric Detection and Minimal Sample Preparation. *J neurochemistry.* 1986;46:1865–76.
51. Huang X, Sterling NW, Du G, Sun D, Stetter C, Kong L, et al. Brain cholesterol metabolism and Parkinson's disease. *Mov Disord.* 2019;34:386–95.
52. Rozani V, Gurevich T, Giladi N, El-Ad B, Tsamir J, Hemo B, et al. Higher serum cholesterol and decreased Parkinson's disease risk: A statin-free cohort study. *Mov Disord.* 2018;33:1298–305.
53. Choi Y, Chan AP. PROVEAN web server: A tool to predict the functional effect of amino acid substitutions and indels. *BIOINFORMATICS.* 2015;31:2745–7.
54. Choi Y, Sims GE, Murphy S, Miller JR, Chan AP. Predicting the functional effect of amino acid substitutions and indels. *PLoS One.* 2012;7:e46688.
55. Lamiable A, Thévenet P, Rey J, Vavrusa M, Derreumaux P, Tufféry P. PEP-FOLD3: faster de novo structure prediction for linear peptides in solution and in complex. *Nucleic acids Res.* 2016;44:W449–W454.
56. Armenteros JJA, Salvatore M, Emanuelsson O, Winther O, von Heijne G, Elofsson A, et al. Detecting sequence signals in targeting peptides using deep learning. *Life Sci alliance.* 2019;2:e201900429.
57. Almagro Armenteros JJ, Sønderby CK, Sønderby SK, Nielsen H, Winther O. DeepLoc: prediction of protein subcellular localization using deep learning. *BIOINFORMATICS.* 2017;33:3387–95.
58. Cho S-I, Lee S, Mok YG, Lim K, Lee J, Lee JM, et al. Targeted A-to-G base editing in human mitochondrial DNA with programmable deaminases. *Cell.* 2022;185:1764–e1712.
59. Teo G, Liu G, Zhang J, Nesvizhskii AI, Gingras A-C, Choi H. SAINTexpress: Improvements and additional features in Significance Analysis of Interactome software. *J Proteom.* 2014;100:37–43.
60. Rhee H-W, Zou P, Udeshi ND, Martell JD, Mootha VK, Carr SA, et al. Proteomic Mapping of Mitochondria in Living Cells via Spatially Restricted Enzymatic Tagging. *Sci (Am Assoc Advancement Sci).* 2013;339:1328–31.
61. Laperle AH, Sances S, Yucer N, Dardov VJ, Garcia VJ, Ho R, et al. iPSC modeling of young-onset Parkinson's disease reveals a molecular signature of disease and novel therapeutic candidates. *Nat Med.* 2020;26:289–99.
62. Kang I, Chu CT, Kaufman BA. The mitochondrial transcription factor TFAM in neurodegeneration: emerging evidence and mechanisms. *FEBS Lett.* 2018;592:793–811.
63. Pyle A, Anugraha H, Kurzawa-Akanbi M, Yarnall A, Burn D, Hudson G. Reduced mitochondrial DNA copy number is a biomarker of Parkinson's disease. *Neurobiol Aging.* 2016;38:216.e217–216.e210.
64. Fukasawa Y, Tsuji J, Fu S-C, Tomii K, Horton P, Imai K. MitoFates: Improved Prediction of Mitochondrial Targeting Sequences and Their Cleavage Sites[S]. *Mol Cell Proteom.* 2015;14:1113–26.
65. Baixauli F, Acin-Perez R, Villarroya-Beltri C, Mazzeo C, Nunez-Andrade N, Gabande-Rodriguez E, et al. Mitochondrial Respiration Controls Lysosomal Function during Inflammatory T Cell Responses. *Cell Metab.* 2015;22:485–98.
66. Schapira AHV, Cooper JM, Dexter D, Clark JB, Jenner P, Marsden CD. Mitochondrial Complex I Deficiency in Parkinson's Disease. *J neurochemistry.* 1990;54:823–7.
67. Jakowec MW, Nixon K, Hogg E, McNeill T, Petzinger GM. Tyrosine hydroxylase and dopamine transporter expression following 1-methyl-4-phenyl-1,2,3,6-tetrahydropyridine-induced neurodegeneration of the mouse nigrostriatal pathway. *J Neurosci Res.* 2004;76:539–50.
68. Hu Z, Mao C, Wang H, Zhang Z, Zhang S, Luo H, et al. CHIP protects against MPP<sup>+</sup>/MPTP-induced damage by regulating Drp1 in two models of Parkinson's disease. *Aging (Albany, NY).* 2021;13:1458–72.
69. Mims MP, Hayes TG, Zheng S, Leal SM, Frolov A, Ittmann MM, et al. Mitochondrial DNA G10398A polymorphism and invasive breast cancer in African-American women [1]. *Cancer Res (Chic, Ill).* 2006;66:1880–1880.
70. Lu J, Wang K, Rodova M, Esteves R, Berry D, Lezi EL, et al. Polymorphic variation in cytochrome oxidase subunit genes. *J Alzheimer's Dis.* 2010;21:141–54.

71. Brakedal B, Dölle C, Riemer F, Ma Y, Nido GS, Skeie GO, et al. The NADPARK study: A randomized phase I trial of nicotinamide riboside supplementation in Parkinson's disease. *Cell Metab.* 2022;34:396–407.e396.
72. Cosi C, Marien M. Decreases in mouse brain NAD<sup>+</sup> and ATP induced by 1-methyl-4-phenyl-1,2,3,6-tetrahydropyridine (MPTP): prevention by the poly(ADP-ribose) polymerase inhibitor, benzamide. *Brain Res.* 1998;809:58–67.
73. Schöndorf DC, Ivanyuk D, Baden P, Sanchez-Martinez A, De Cicco S, Yu C, et al. The NAD<sup>+</sup> Precursor Nicotinamide Riboside Rescues Mitochondrial Defects and Neuronal Loss in iPSC and Fly Models of Parkinson's Disease. *Cell Rep. (Camb).* 2018;23:2976–88.
74. Schwab AJ, Sison SL, Meade MR, Broniowska KA, Corbett JA, Ebert AD. Decreased Sirtuin Deacetylase Activity in LRRK2 G2019S iPSC-Derived Dopaminergic Neurons. *Stem cell Rep.* 2017;9:1839–52.
75. Wakade C, Chong R, Bradley E, Thomas B, Morgan J. Upregulation of GPR109A in Parkinson's disease. *PLoS one.* 2014;9:e109818–e109818.
76. Dellinger RW, Santos SR, Morris M, Evans M, Alminana D, Guarente L, et al. Repeat dose NRPT (nicotinamide riboside and pterostilbene) increases NAD<sup>+</sup> levels in humans safely and sustainably: a randomized, double-blind, placebo-controlled study. *npj aging mechanisms Dis.* 2017;3:17–19.
77. Straccia G, Colucci F, Eleopra R, Cilia R. Precision Medicine in Parkinson's Disease: From Genetic Risk Signals to Personalized Therapy. *Brain Sci.* 2022;12:1308.

## ACKNOWLEDGEMENTS

This work was supported by Department of Defense (W81XWH2110625) to SJK and by NIH (P01AG034906, R01AG068405, P30AG068345) to PC. Targeted NAD<sup>+</sup> Metabolomics was supported by the University of Washington Nathan Shock Center of Excellence in the Basic Biology of Aging (3 P30 AG013280). HeLa expressing HA-EGFP-OMP25 (HA-MITO) or myc-EGFP-OMP25 (CON-MITO) cells were gifts from Dr. Sabatini. Diagram of the mitochondrial DNA was provided by Ana Silverstein.

## AUTHOR CONTRIBUTIONS

SK created the framework of the study, wrote the paper, and performed a majority of the experiments. BM, NH, NAG performed mass spec analysis. BM and KY wrote the paper. RR performed stability and NAD experiment. RB performed APEX experiment. NL and YW performed mitochondria isolation. AH performed TFAM knockdown experiment. TEA and EMC contributed to HRS data analysis and PW, XS, CL, and DL contributed to FHS data analysis. GMP and MWJ performed MPTP mice experiment. PC supervised the project.

## FUNDING

Open access funding provided by SCEL, Statewide California Electronic Library Consortium.

## COMPETING INTERESTS

Pinchas Cohen is a consultant of CohBar Inc.

## ADDITIONAL INFORMATION

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41380-023-02344-0>.

**Correspondence** and requests for materials should be addressed to Pinchas Cohen.

**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 licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence 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 licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023