Nucleic Acids Research, 2025, **53**, gkaf1072
https://doi.org/10.1093/nar/gkaf1072
Molecular and Structural Biology

Pooled overexpression screening identifies PIPPI as a novel microprotein involved in the ER stress response

Lorenzo Lafranchi[1,2], Glancis Luzeena Raja[1,2], Alberto M. Arenas[1,2], Anna Spinner[1,2], Kovi R. Shrung[3], Maximilian Hornisch[4], Dörte Schlesinger[1,2], Carmen Navarro Luzon[1,2], Linus Brinkenstråhle[1,2], Rui Shao[1,2], Ilaria Piazza[4], Janne Lehtiö[3], Rui M.M. Branca[3], Simon J. Elsässer[1,2,*]

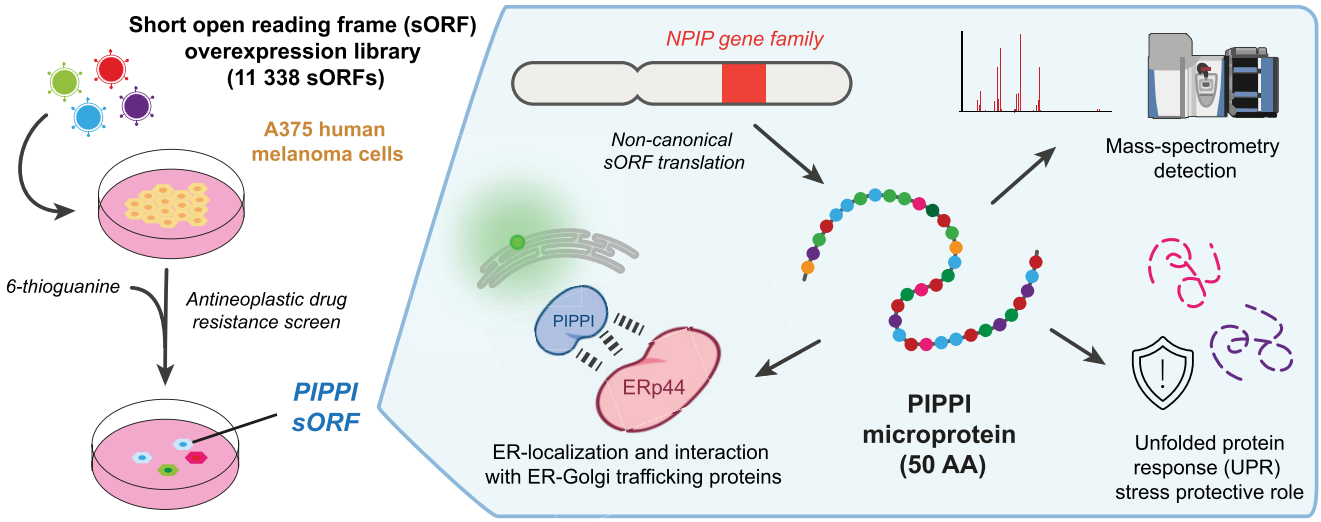
[1]Science for Life Laboratory, Karolinska Institutet, Department of Medical Biochemistry and Biophysics, Division of Genome Biology, Solna 17165, Sweden
[2]Ming Wai Lau Centre for Reparative Medicine, Stockholm node, Karolinska Institutet, Solna 17165, Sweden
[3]Science for Life Laboratory, Karolinska Institutet, Department of Oncology-Pathology, Solna 17165, Sweden
[4]Max Delbrück Center for Molecular Medicine in the Helmholtz Association (MDC Berlin), Berlin 13125, Germany

*To whom correspondence should be addressed. Email: simon.elsasser@scilifelab.se

Abstract

Microproteins encoded by short open reading frames (sORFs) of <100 codons have been predicted to constitute a substantial fraction of the eukaryotic proteome. However, the relevance and roles of a majority of microproteins remain undefined, as only a small fraction of these intriguing cellular players have been characterized in-depth so far. Here, we use pooled overexpression screens with a library of 11 338 sORFs to overcome the challenge of elucidating which of the thousands of putative translated sORFs are biologically functional. As a proof-of-concept, we performed a phenotypic screen to identify sORFs protecting cells from treatment with the nucleotide analogue 6-thioguanine. With this approach, we identified two cytoprotective microproteins: altDDIT3 and PIPPI. PIPPI is encoded within the LCR16a core duplicon of the Morpheus/NPIP gene cluster. We show that PIPPI modulates the cellular response to protein folding stress in the endoplasmic reticulum (ER) and interacts with proteins in the same pathway, including protein disulfide isomerase ERp44. PIPPI overexpression protects, while PIPPI knockdown sensitizes cells to ER stress. Besides providing mechanistic insights into a new microprotein, this study highlights the power of using pooled overexpression screens to identify functional microproteins.

Graphical abstract



Received: December 8, 2024. **Revised:** September 21, 2025. **Accepted:** September 25, 2025

Introduction

Short open reading frames (sORFs) with the potential to encode functional polypeptides, have remained unnoticed for a long time due to their small size (i.e. <300 bp). Nevertheless, increasing evidence highlights that these microproteins, or micropeptides, are abundantly expressed in both prokaryotic and eukaryotic cells. sORFs can be present in messenger RNA (mRNA) molecules, as an alternative to the main ORF, but also in other transcripts, such as long noncoding RNAs, microRNAs, or antisense transcripts [1, 2]. Over recent years, a significant effort has been undertaken in distinguishing between noncoding and translating sORFs, for which ribosome profiling (Ribo-seq), mass spectrometry (MS), and bioinformatic assessment of evolutionary conservation, have served as the main approaches used for identifying putative microproteins [1]. Finally, ongoing efforts are aiming to benchmark the ever-growing number of sORFs predicted to be translated using Ribo-seq, and propose a high-confidence and standardized catalogue of human sORFs [3]. Although thousands of microproteins have been predicted to exist in mammalian cells, only a small fraction of them have been thoroughly characterized. To mention a few prominent examples, microproteins have been shown to be involved in the maintenance of genome stability [4–6], gene expression [7, 8], metabolism [9, 10], and signalling [11–14]. Even from this brief overview, it can be appreciated that microproteins display an intriguing variety of target tissues, subcellular compartments, and intra- or extracellular mechanisms of action [15]. Ongoing research is highlighting how microproteins are not specifically evolved to solely serve in a subset of biological processes, but they appear to play an underappreciated role in cellular and organismal biology [15, 16].

Cells are equipped with a plethora of evolutionarily conserved signalling pathways that can be activated to counteract endogenous and exogenous stress events, thereby ensuring the maintenance of cellular homeostasis. In case of excessive stress, pro-death pathways represent a last resort used to remove severely damaged cells from the organism [17]. Recent ribosome profiling studies have demonstrated that translation of a large number of sORFs, specifically upstream ORFs (uORFs), is upregulated in stress conditions [18–21]. This is particularly relevant for cancer initiation and progression, as tumor cells exhibit remarkable plasticity and the ability to rewire their intracellular pathways to overcome negative stimuli, such as antineoplastic drug treatments, which ultimately leads to drug resistance in cancer patients [22]. While some of these translational events could lack function at the protein level, we can expect that many functional stress-induced microproteins await discovery and characterization [23], potentially posing themselves as clinical biomarkers or molecular targets for cancer and other diseases [24]. To date, two of the major challenges that remain in the microprotein field are to elucidate which of the thousands of putative translated sORFs are biologically active, and to functionally characterize them. Here, we sought to identify bioactive microproteins with a pooled overexpression screen of putative sORFs encoded in the human genome. Upon introducing an sORF library into human A375 melanoma cells, we selected for increased resistance to the widely used chemotherapeutic compound 6-thioguanine (6-TG), identified two sORFs promoting cell proliferation, and followed up with the in-depth characterization of one of the candidates.

Materials and methods

Cell culture

A375, U2OS, HEK293T, and COS-7 cell lines were maintained in Dulbecco's modified Eagle medium (DMEM), containing high glucose, GlutaMAX™ and pyruvate (Gibco). hTERT RPE-1 cells were maintained in DMEM/Nutrient Mixture F-12 (DMEM/F-12), containing high glucose, GlutaMAX™ and pyruvate (Gibco). For all cell lines, medium was supplemented with 10% fetal bovine serum (FBS; Sigma–Aldrich). All cell lines were cultured in an ambient-controlled incubator at 37 °C, 5% O_2, and 5% CO_2. A375 and COS-7 cells were purchased from Sigma–Aldrich. HEK293T cells were a gift from Prof. T. Helleday, whereas U2OS and RPE-1 cells were a gift from Prof. J. Bartek.

DNA constructs

DNA constructs and sources are detailed in Table1.

Generation of stable expression cell lines using lentivirus

Twenty-four hours post seeding, HEK293T cells were co-transfected with a lentiviral transfer plasmid, the envelope plasmid pMD2.G (Addgene #12260), and the second generation packaging plasmid psPAX2 (Addgene #12259), using transIT®-LT1 (Mirus Bio) according to manufacturer's instructions. After 24 h, medium was replenished with fresh DMEM. Virus supernatant was collected at 48 and 72 h and, after filtration using a 0.45 mm mixed cellulose esters syringe filter (Millipore), directly used to transduce target cells. The following concentrations of polybrene (Sigma–Aldrich), puromycin (Cayman Chemical), or blasticidin (InvivoGen) were used to either help infect or select target cells: A375 (2 μg/ml, 3 μg/ml, 20 μg/ml), U2OS (4 μg/ml, 6 μg/ml, -), hTERT RPE-1 (4 μg/ml, -, 50 μg/ml).

sORF library design and production

sORF sequences composing the library (Supplementary Table S1) were ordered at Twist Biosciences as a 200-nucleotide oligo pool. Because of this hard limit imposed by DNA synthesis, only microproteins up to 57 amino acids (AA) in length were included in the library. Coding sequences have been codon optimized for human expression using Twist's internal algorithm. An ATG starting codon was included for all coding sequences and Ochre (TAA) was selected to be the consensus stop codon. SapI sites surrounded the sORF sequences to enable cloning into the modified recipient pCDH_EF1_IRES_Puro lentiviral expression plasmid (System Biosciences, CD532A-2). Following golden gate cloning using SapI (NEB) and T4 ligase (NEB), the vector was electroporated into MegaX DH10BTM $T1^R$ cells (Invitrogen). Bacteria were plated and selected overnight with 200 mg/ml carbenicillin. At this stage, multiple plates ensuring the maintenance of library diversity were used. Colonies were collected by scraping, and pooled before extracting plasmid DNA using the QIAGEN Plasmid Plus Maxi Kit (Qiagen).

Virus production for pooled screens

A concentrated lentivirus pool was produced at the VirusTech Core Facility of Karolinska Institutet. In short, HEK293T cells were co-transfected with the lentiviral expression plasmid pCDH_EF1-sORFs_IRES_Puro, the envelope plasmid

Table 1. DNA constructs

ID	Name/description	Backbone	Origin
CD532A-2	pCDH_EF1a_MCS_IRES_Puro	pCDH	System Biosciences
L300	pCDH_EF1a_MCS_IRES_Puro_SapI_removed	CD532-A	This study
L301	pCDH_EF1a_2xSapI_A_IRES_Puro	L300	This study
L302	pCDH_EF1a_2x SapI_C_IRES_Puro	L300	This study
L303	pCDH_EF1a_2x SapI_G_IRES_Puro	L300	This study
L309	pCDH_EF1a_All-sORFs_IRES_Puro	L301/L302/L303	This study
L320	pCDH_EF1a_altDDIT3_IRES_Puro	L302	This study
#12260	psPAX2 (second generation lentiviral packaging plasmid)		Addgene
#12259	pMD2.G (VSV-G envelope expressing plasmid)		Addgene
L321	pCDH_EF1a_altDDIT3-GFP_IRES_Puro	L327	This study
L325	pAS1_EF1a_PIPPI-HA_IRES_Puro	E404 (#140009)	This study
K106	pAS1_4 × 7SKPylT_EF1_Ub-∗PIPPI	K038 (#162801)	This study
K110	pAS1_4 × 7SKPylT_EF1_Ub-∗PIPPI-HA	K039 (#162802)	This study
K111	pAS1_4 × 7SKPylT_EF1_Ub-PIPPI-HA	K039 (#162802)	This study
L327	pCDH_EF1a_2x SapI_C-GFP_IRES_Puro	L302	This study
L329	pCDH_EF1a_PIPPI_IRES_Puro	L302	This study
L356	pCDH_EF1a_GFP_IRES_Puro	L313	This study
L330	pCDH_EF1a_PIPPI-GFP_IRES_Puro	L327	This study
L344	pCDH_EF1a_PIPPI-L-GFP_IRES_Puro	L327	This study
L346	pCDH_EF1a_2x SapI_C-GFP_IRES_BSD	L327	This study
L348	pCDH_EF1a_PIPPI-GFP_IRES_BSD	L346	This study
#115969	pLVX-ATF4 mScarlet NLS	pLVX-Puro	Addgene
L363	pCDH_EF1a_PIPPI-L_IRES_Puro	L302	This study
L381	pCDH_EF1a_PIPPI-del2-9-GFP_IRES_Puro	L327	This study
L382	pCDH_EF1a_PIPPI-del10-17-GFP_IRES_Puro	L327	This study
L383	pCDH_EF1a_PIPPI-del20-27-GFP_IRES_Puro	L327	This study
L384	pCDH_EF1a_PIPPI-del28-35-GFP_IRES_Puro	L327	This study
L385	pCDH_EF1a_PIPPI-del36-43-GFP_IRES_Puro	L327	This study
L386	pCDH_EF1a_PIPPI-del44-51-GFP_IRES_Puro	L327	This study
A014	pCDH_EF1a_2x SapI_G_IRES_BSD	L303	This study
L388	pCDH_EF1a_HA-ERp44_IRES_BSD	A014	This study
L389	pCDH_EF1a_HA-ERp44_CRFS-mut_IRES_BSD	A014	This study
L405	pCDH_EF1a_PIPPI-C35A_IRES_Puro	L302	This study
L406	pCDH_EF1a_PIPPI-C35A-GFP_IRES_Puro	L327	This study
L407	pCDH_EF1a_PIPPI-C51A_IRES_Puro	L302	This study
L408	pCDH_EF1a_PIPPI-C51A-GFP_IRES_Puro	L327	This study
L410	pCDH_EF1a_PIPPI-Q12N-GFP_IRES_Puro	L327	This study

pMD2.G (Addgene #12259), and the second generation packaging plasmid psPAX2 (Addgene #12260). pMD2.G and psPAX2 plasmids were gifts from Didier Trono. The media was exchanged after 16 h, and virus-containing supernatant was collected 48 h and 64 h post transfection. After concentrating by double step centrifugation, the final titre was assessed by extracting and quantifying proviral DNA from transduced cells. The number of integrations into the host genome was calculated by normalizing the total number of provirus copies to a housekeeping gene (hALB).

sORF screens

During screens, media was supplemented with 1% penicillin–streptomycin. A375 cells were transduced at day 0 with the lentiviral sORF-encoding library in two biological replicates at a low multiplicity of infection (MOI) (∼0.5). For each replicate, transduction was performed in presence of 2 μg/ml polybrene (Sigma–Aldrich), and in enough cells to achieve a representation of at least 1000 cells per sORF. Two days after transduction, cells were re-seeded and selected with 1.5 mg/ml puromycin. At day 5, a fraction of cells was pelleted by centrifugation and frozen for genomic DNA (gDNA) extraction (timepoint 0, "T0"). The remaining cells were kept in puromycin and split in different arms for further "no treatment" or drug treatments (see Fig. 1). After one day, puromycin selection was removed, and treatment with 10 μM 6-TG (Sigma–Aldrich) was started. Cells of the "no treatment" control arm were maintained under puromycin selection for three days longer. Untreated cells were sub-cultured every 3–4 days until the final timepoint at day 21. Drug treatment was renewed every third day, and cells were subcultured when confluent. The screen was continued until cells repopulated the culture flasks (day 28). Throughout the screens, cells were collected or re-seeded in a minimal number that would maintain a representation of at least 1000 cells per sORF.

gDNA extraction and sequencing

gDNA was isolated using the QIAamp DNA Blood Midi Kit (Qiagen) according to the manufacturer's instructions. For PCR amplification of sORF sequences, gDNA was divided into 100 μl reactions such that each tube had at most 4 μg of gDNA. PCR reactions were performed using NEBNext® Ultra™ II Q5® Master Mix (NEB). Afterwards, tubes were pooled per sample and purified using the QIAquick PCR Purification Kit (Qiagen). Custom-made primers containing the i5/i7 adapter sequences were synthesized at Integrated DNA Technologies (IDT), and incorporated during a second PCR step. PCR products were purified using QIAquick PCR Purification Kit (Qiagen), and sequenced on an Illumina NextSeq500 platform. Finally, the resulting Fastq files were aligned to our custom sORF library using kallisto v0.46.1 [25].

Table 2. PIPPI-targetingsiRNAs

siRNA	5′–3′ RNA sequence
PIPPI_siRNA_1_sense	GUUGACCUGGUGCAAGAAACAUGCT
PIPPI_siRNA_1_antisense	AGCAUGUUUCUUGCACCAGGUCAACAG
PIPPI_siRNA_2_sense	CUGAGAAUCAGAGACAGAACAUAAA
PIPPI_siRNA_2_antisense	UUUAUGUUCUGUCUCUGAUUCUCAGAA
PIPPI_siRNA_3_sense	ACAUUCUCCAAAAAGAAGCUGCAAC
PIPPI_siRNA_3_antisense	GUUGCAGCUUCUUUUUGGAGAAUGUUU
PIPPI_siRNA_4_sense	GGAGCAGGGUGGUCACACUGACCTT
PIPPI_siRNA_4_antisense	AAGGUCAGUGUGACCACCCUGCUCCAA

Growth competition assays and flow cytometry

One day before treatment, cell lines were seeded in a 1:1 ratio. This time point is considered "Day 0" of each assay, and an aliquot of the cell mix was collected for flow cytometry analysis. On day 1, cells were treated with the indicated dose of 6-TG (Sigma–Aldrich) or tunicamycin (TM; Sigma–Aldrich) (see figures for details). An untreated control well was included for each replica, and biological replicas were performed on different days. Throughout the duration of the assay, cells were visually assessed for viability and confluency, and split accordingly. On sampling days, cells were collected by trypsinization, and a fraction of the suspension was fixed with 4% formaldehyde for 10 min at room temperature. After washing, cells were resuspended in phosphate buffered saline (PBS), and stored in the dark at 4 °C until analysing them with a Navios flow cytometer (Beckman Coulter). At least 10 000 living cells were measured for each sample. Long-term growth competition assays shown in Supplementary Fig. S1B were performed as above, but 6-TG was renewed every third day, and cells were collected for flow cytometry every seventh day. In this case, biological replicates were performed in parallel.

For siRNA knockdown and competition assays, A375 parental and A375-GFP cells were used. Four different siRNAs were designed using an online siRNA design tool (https://eu.idtdna.com/site/order/tool/index/DSIRNA_PREDESIGN), and ordered from IDT along with a nontargeting, negative control (NC) siRNA (IDT, #51-01-14-03). The sequences of all siRNAs are listed in Table 2. Cells were transfected with either a nontargeting control (NC) siRNA, or a pool of siRNAs (10 μM) targeting PIPPI, using lipofectamine RNAiMAX (Thermo Scientific, #13778075) according to the manufacturers' recommendations. Twenty-four hours after transfection, three different 1:1 co-cultures of A375 cells transfected with PIPPI siRNAs were established: a control co-culture (Parental + NC: GFP + NC), a PIPPI-knockdown GFP-negative co-culture (Parental + siRNA: GFP + NC), and a PIPPI-knockdown GFP-positive co-culture (Parental + NC: GFP + siRNA). The initial percentage of GFP-positive/negative cells was quantified by flow cytometry using an Attune TxT cytometer (Thermo Scientific) the day the co-cultures were plated. Then, cells were maintained in culture for 7 days, either with vehicle dimethylsulfoxid (DMSO) or TM (500 ng/ml) treatment, with passages and renewal of drug every 2–3 days to prevent full confluency. Final GFP-positive/negative percentages were measured again by flow-cytometry on day 7. The flow cytometry results were analysed using FlowJo™ v10.8 Software (BD Life Sciences).

Generation of PIPPI antibody

The PIPPI antibody was generated by immunoGlobe GmbH. In short, a rabbit was immunized using the 11-AA-long synthetic peptide CSENQRQNIKG, corresponding to PIPPI AA 14–23. Serum from the final bleed was affinity-purified using the High-Affinity Antibody Purification Kit (GenScript) with the immobilized synthetic peptide, and used for immunoblotting (1:200–1:500) or immunoprecipitation experiments.

Protein gels and western blotting

Cell lysates were prepared in RIPA buffer [50 mM Tris–HCl, pH 7.5, 1% NP-40, 0.25% sodium deoxycholate, 150 mM NaCl, 1 mM ethylenediaminetetraacetic acid (EDTA), and 0.1% sodium dodecyl sulphate (SDS)] supplemented with 1× cOmplete™ Protease inhibitor cocktail (Roche). After sonication, the insoluble fraction was removed by centrifugation. Proteins were resolved by sodium dodecyl sulphate–polyacrylamide gel electrophoresis (SDS–PAGE) under denaturing conditions using 4%–20% Mini-PROTEAN® TGX™ precast polyacrylamide gels (Bio-Rad). Proteins were transferred to nitrocellulose membranes (Bio-Rad). Immunoblots were performed using the appropriate primary antibodies (Table 3) and the relative horseradish peroxidase (HRP)-coupled secondary antibodies (Bio-Rad). Proteins were visualized on an ImageQuant™ LAS500 Imager (GE Healthcare Life Sciences) using the Immobilon Forte Western HRP substrate (Merck Millipore).

To visualize ncAA(TCO∗K)-containing proteins, 500 nM SiR-tetrazine (Spirochrome) was added to the lysates for 15 min at room temperature before denaturation. After electrophoresis, the gel was imaged using an Amersham Imager 600 (GE Healthcare life Sciences).

Amber suppression

HEK293T were co-transfected with a plasmid encoding the protein of interest, and a plasmid encoding the *AF* variant of the *Methanosarcina mazei* pyrrolysine transfer RNA synthetase (PylRS-AF), using transIT®-LT1 (Mirus Bio) according to manufacturer's instructions. COS-7 cells were transfected using Lipofectamine LTX™ with PLUS™ reagent (Invitrogen) according to manufacturer's protocol. Axial trans-cyclooct-2-ene-l-lysine (TCO∗K) was added at the time of transfection, and cells were either fixed after 24 h for microscopy (COS-7), or lysed after 48 h and analysed by immunoblotting (HEK293T). When cells were to be analysed by microscopy, the TCO∗K was withdrawn 1 h before cell fixation. TCO∗K (SiChem, SC-8008) stock solution was prepared at 100 mM in 0.2 M NaOH/H_2O, 15% DMSO, and diluted to 50 μM right before use in the appropriate growth medium. At the end of the experiment, TCO∗K-containing polypeptides were labeled using either Tetrazine-Silicon Rhodamine (tet-SiR, Spirochrome) or 6-Methyl-Tetrazine-BODIPY®-FL (me-tet-BDP-FL, Jena Bioscience). Both stocks were prepared in dimethylformamide (DMF) and further diluted in either RIPA buffer (lysate labeling) or Tris-buffered saline with Tween 20 (TBS-T) for fixed cells labeling before use.

Analysis of glycosylated proteins

Constructs were forward transfected in HEK293T cells using transIT®-LT1 (Mirus Bio) according to manufacturer's instructions. After 24 h, cells were lysed in RIPA buffer and, after

Table 3. Antibodies used in the study

Target	Antibody name	Company	Application
ATF-4	Anti-ATF-4 (D4B8) rabbit mAb, #11815	Cell Signaling	WB (1:1000)
β-Actin	β-Actin (13E5) rabbit mAb, #4970	Cell Signaling	WB (1:5000)
BiP	BiP (C50B12) rabbit mAb, #3177	Cell Signaling	WB (1:1000)
Calnexin	Calnexin (C5C9) rabbit mAb, #2679	Cell Signaling	WB (1:1000), IF (1:50)
CHOP	CHOP (L63F7) mouse mAb, #2895	Cell Signaling	WB (1:500)
ERp44	ERp44 rabbit Ab, #2886	Cell Signaling	WB (1:500)
ERp44	TXNDC4 (E-6) mouse mAb, sc-393687	Santa Cruz	IP
GAPDH	GAPDH chicken Ab, AB2302	Millipore	WB (1:10 000)
GFP	GFP (B-2) mouse mAb, sc-9996	Santa Cruz	WB (1:5000)
GFP	GFP rabbit Ab, pabg-1	ChromoTek	IF (1:800)
HA	HA-HRP high affinity Ab, 12013819001	Roche	WB (1:5000)
HA	HA-Tag (F-7) mouse mAb, sc-7392w	Santa Cruz	IF (1:400)
P4HB	PDI (C81H6) rabbit mAb, #3501	Cell Signaling	WB (1:1000)
SLC30A7	SLC30A7 rabbit Ab, HPA018034	Atlas Antibodies	WB (1:200)
Vinculin	Vinculin, rabbit Ab, #4650	Cell Signaling	WB (1:1000)
PIPPI	Polyclonal rabbit serum	immunoGlobe GmbH	WB (1:200–1:500)

sonication, the insoluble fraction was removed by centrifugation. Samples were then treated with PNGase F (NEB) according to manufacturer's instructions. In short, lysates were denatured for 10 min in Denaturing Buffer at 100°C. After cooling, samples were incubated for 60 min at 37°C in the presence of 1% NP-40, GlycoBuffer 2, and PNGase F. NC samples were included. All samples were then analysed by SDS–PAGE gels and immunoblotting.

Immunofluorescence microscopy

Cells grown in 18 well μ-Slides (ibidi) were fixed with 4% formaldehyde for 10 min at room temperature and permeabilized for 15 min with 0.1% (v/v) triton/PBS. After washing with TBS supplemented with 0.1% Tween-20 (TBS-T), cells were blocked with 2% BSA in TBS-T for 1 h and then incubated overnight at 4 °C with primary antibodies (Table3). In amber suppression experiments, cells were click-labeled for 60 min with 500 nM 6-Methyl-tetrazine-BODIPY-FL in blocking buffer (2% BSA in TBS-T) and washed three times with TBS-T prior to the incubation with primary antibodies. After washing with TBS-T, cells were stained with Alexa Fluor-conjugated secondary antibodies (Life Technologies) for 60 min at room temperature and counterstained with 1 mg/ml 4′,6-diamidino-2-phenylindole DAPI (Sigma–Aldrich) or with 10 μg/ml AF647-conjugated Concanavalin A (Con A; Invitrogen). After washing, cells were imaged on a Zeiss LSM880 confocal laser scanning microscope using a 63×/1.4 oil immersion objective. Images were processed and prepared for publication using Fiji [26].

Live cell microscopy

One day before the experiment, A375 cells expressing the ATF4-mScarlet-NLS reporter were seeded in 18 well μ-Slides (ibidi) and cultured overnight in DMEM supplemented as detailed above. One hour before starting the experiment, medium was exchanged to Leibovitz's L-15 Medium (no phenol red, Gibco) supplemented with 10% FBS (Sigma–Aldrich), 4.5 g/l glucose (Sigma–Aldrich), and 1 mM sodium pyruvate (Gibco). Prior to imaging, TM (500 ng/ml) was added to some of the wells. Cells were imaged over 24 h at 2-h intervals using a Nikon eclipse Ti2 inverted widefield microscope equipped with a heated imaging chamber. Images were acquired with a 20×/0.75 air objective and, post-acquisition, quantified using CellProfiler [27]. Plots included in the publications were prepared using ggplot2 v3.3.5 [28] run on RStudio v2022.02.03 [29].

Co-immunoprecipitation/MS

Cell lysates were prepared from 80% confluent T-175 flasks in RIPA buffer not containing SDS (50 mM Tris–HCl, pH 7.5, 1% NP-40, 0.25% sodium deoxycholate, 150 mM NaCl, 1 mM EDTA) supplemented with 1× cOmplete™ Protease inhibitor cocktail (Roche). After allowing lysis for 15 min, rotating at 4 °C, the insoluble fraction was removed by centrifugation. Supernatants were incubated for 4 h rotating at 4 °C with either 25 μl of GFP-trap magnetic beads (ChromTek, pre-equilibrated in RIPA w/o SDS) or anti-HA magnetic beads (Pierce™, pre-equilibrated in 0.05% TBS-T). In the case of the IP-LC MS/MS experiment in U2OS cells, RIPA was supplemented with 0.1% SDS for lysis. Before incubating the lysates with the magnetic beads, samples were diluted 1:10 in RIPA w/o SDS to lower the concentration of the detergent. After incubation with the lysates, beads were bound to a magnetic stand and, if not specified otherwise, washed three times with RIPA w/o SDS and once with ddH_2O. Samples were shortly vortexed in-between washes. Finally, proteins were eluted from beads either by boiling in 2× Laemmli buffer (4% SDS, 20% glycerol, 120 mM, pH 6.8) for western blotting analysis or using 1% acetic acid for liquid chromatography-tandem mass spectrometry (LC-MS/MS analysis). Elution with 1% acetic acid was performed twice and supernatants were pooled and evaporated by heating.

The dried samples were reduced and alkylated for 30 min at room temperature with 5 mM Tris(2-carboxyethyl)phosphine TCEP (Pierce™) and 20 mM chloroacetamide (Sigma) dissolved in 250 mM Tris buffer, pH 8.5. Sera-Mag beads (Sigma–Aldrich) were added to the samples in a 1:25 ratio. An equal volume of absolute ethanol was added to the mixture to stimulate coupling of the proteins to beads. After 5 min incubation at room temperature, beads were washed on a magnetic stand three times with 80% ethanol and then resuspended in 50 mM triethylammonium bicarbonate TEAB (Sigma–Aldrich).One microgram of trypsin (Pierce™) was added to each sample and digestion was allowed to occur overnight at 37 °C under gentle agitation. After immobilizing beads on a magnetic stand, supernatants were collected. Bound peptides were then eluted using 2% acetoni-

trile (ACN; Sigma–Aldrich) in 50 mM TEAB (Sigma–Aldrich) and pooled with the collected supernatants. Samples were completely dried using a SpeedVac and finally resuspended in solvent B (80% ACN, 0.1% formaldehyde) for LC-MS/MS analysis.

LC-MS/MS service was provided by the Finkemeier lab (Münster University). Samples were analysed using an EASY-nLC 1200 (Thermo Fisher Scientific) coupled to an Exploris 480 mass spectrometer (Thermo Fisher Scientific). Separation of peptides was performed on 20 cm frit-less silica emitters (CoAnn Technologies, 0.75 µm inner diameter), packed in-house with reversed-phase ReproSil-Pur C18 AQ 1.9 µm resin (Dr Maisch). The column was constantly kept at 50 °C. Peptides were eluted in 115 min applying a segmented linear gradient of 0%–98% solvent B [solvent A 0% ACN, 0.1% FA; solvent B 80% ACN, 0.1% formic acid (FA)] at a flowrate of 300 nl/min. Mass spectra were acquired in data-dependent acquisition mode. MS1 scans were acquired at an Orbitrap Resolution of 120 000 with a Scan Range (m/z) of 380–1500, a maximum injection time of 100 ms and a Normalized AGC Target of 300%. For fragmentation only precursors with charge states 2–6 were considered. Up to 20 Dependent Scans were taken. For dynamic exclusion, the exclusion duration was set to 40 s and a mass tolerance of ± 10 ppm. The Isolation Window was set to 1.6 m/z with no offset. A normalized collision energy of 30 was used. MS2 scans were taken at an Orbitrap Resolution of 15 000, with a fixed First Mass (m/z) = 120. Maximum injection time was 22 ms and the normalized AGC Target 50%.

The raw data were analysed using MaxQuant Version 1.6.3.4 [30] and searched against the Uniprot protein database (Human all 2017/11) [31], PIPPI sequences, and the CRAPome database [32]. MaxQuant default settings were used with the two options *Match between runs* and *LFQ intensity reporting* activated. The resulting LFQ intensities were then analysed using the Differential Enrichment analysis of Proteomics data (DEP) package 1.14.0 [33] on RStudio 2022.02.03. After removing common contaminants, imputation was performed using the "MinProb" method without a prior "DEP normalization" of the data. Proteins were then ranked according to their LFC enrichment over parental control and DEP *P*-value. Only proteins present in 6/6 replicates of each pulldown were considered for this analysis. Finally, volcano plots for publication were generated using the VolcaNoseR web app [34].

PIPPI-enrichment/MS

For immunoprecipitation/MS experiments with the PIPPI rabbit antibody in Fig. 2 and Supplementary Fig. S2, Pierce protein-A magnetic beads (50 µl/reaction) were washed twice with 1× PBS and incubated in the PIPPI rabbit antibody (1:100 dilution in 1× PBS) overnight at 4 °C.The following day, antibody-bound beads were washed once with RIPA buffer to equilibrate them. 2 Mio A375 cells (parental, PIPPI, and PIPPI-GFP) were lysed in RIPA buffer for 10 min on ice, followed by centrifugation at 16 000 ×*g* at 4 °C for 15 min. The supernatant lysates were retained andadded to antibody-bound beads. The bead + lysate suspension was incubated overnight at 4 °C. Post incubation, the beads were washed once in RIPA buffer, once in 1× PBS with 250 mM NaCl, twice with 1× PBS, and once with ddH_2O. The beads were then eluted twice with 1% acetic acid, incubated for 5 min each. The eluates from both rounds were pooled and then evaporated at 95 °C for 25 min, followed by reduction/alkylation with 5 mM TCEP (Sigma–Aldrich) and then directly alkylated in the presence of 20 mM 2-chloroacetamide (Sigma–Aldrich) for 30 min at 22 °C. Next, proteins were pre-digested for 2 h at 37 °C with 1 µg Lys-C (Fujifilm Wako Chemicals), followed by digestion with trypsin at an enzyme: substrate ratio of 1:100 (Promega), and incubation overnight at 37 °C. The following day, digests were adjusted with 0.1% trifluoroacetate TFA for cleanup using C18 tips (Thermo) per manufacturer instructions. The cleared peptides were dried in a SpeedVac for 2 h at 65 °C. For MS measurement, peptides were reconstituted in 0.1% FA.

For the endogenous detection of PIPPI peptides in A375 parental, PIPPI, and PIPPI-GFP cells (Fig. 2D and E and Supplementary Fig. S2H), we used an Orbitrap Astral and timsTOF HT-Evosep mass spectrometer, for which protocols of sample processing are described below:

Astral

The following protocol was used to detect PIPPI peptides in IP/MS samples with a data dependent acquisition (DDA) in Fig. 2E: top left – Peptide 1, and top right – Peptide 2 MS2 spectra, Supplementary Fig. S2H: top left – Peptide 1, top two right – Peptide 2 MS2 spectra. The VanquishNeo UPLC system (Thermo Scientific) injected samples (containing 200 ng of digested peptides dissolved in 0.1% FA) into the trap column (PepMap™ Neo Trap Cartridge, Thermo Scientific) with a 150 µl loading volume. Then followed a 41.5-min separation gradient (mobile phase A: 0.1% FA in H_2O, and mobile phase B: 80% ACN, 0.1% FA) on an Aurora Ultimate G3, C18, 25 cm × 75 µm ID, 1.7 µm bead size analytical column (IonOpticks, Melbourne, Australia) online with nanoelectrospray onto the Orbitrap Astral Mass Spectrometer (Thermo Scientific). The gradient went from 1% to 4% ACN in 0.5 min, to 32% in 25.5 min, to 45% in 30.5 min, and finally to 99% at 32.5 min to initiate a washout for 10 min. For all of these acquisitions, flowrate was maintained at 300 nl/min and the MS instrumentation parameters were as follows: For a full scan in positive mode, Orbitrap resolution was set to 120 000, scan range m/z 380–1180, RF Lens 45%, with a target AGC of 300% or until a maximum injection time of 10 ms is reached. Precursors were selected for MS/MS scans in the Astral mass analyzer based on a Intensity threshold filter of 10^3, and charge states 2, 3, 4, 5, 6 were selected, with a dynamic exclusion duration of 25 s. MS/MS was performed with an isolation window of 2 Th, 27% HCD collision energy, with a scan range of 100–1200 m/z, with a target AGC of 100% or until the maximum injection time of 10 ms is reached.

timsTOF

The following protocol was used to detect PIPPI peptides in IP/MS samples on the timsTOF HT-Evosep with a ddaPASEF-based workflow in Fig. 2E: bottom right – Peptide 2 MS2 spectra,Supplementary Fig. S2H: bottom right – Peptide 2 MS2 spectra. For MS analysis, 200 ng of peptides were loaded onto Evotip PURE (EV2013) according to manufacturer's instructions (EvoSep). The samples were separated using a standard 30SPD gradient (44 min). The separation column was a PepSep FIFTEEN 15 cm × 150 µm × 1.5 µm (Bruker 1893474) connected to a 20 µM ZDV Sprayer (Bruker 1865710) kept at 40 °C by a heating oven (Column Toaster, Bruker, Bremen, Germany). Mobile phase A was 0.1%

FA in MQ and B 0.1% FA in ACN. Online LC-MS was performed using a Tims TOF HT mass spectrometer (Bruker) in DDA mode, using the CaptiveSpray source, capillary voltage 1500 V in positive mode, dry gas flow of 3 l/min, dry gas temperature at 180 °C. Mass range 100–1700 m/z and mobility range of 0.6–1.6 V.s.cm2 was selected. MS/MS was used with 10 PASEF (parallel accumulation—serial fragmentation) scans (with 100 ms ramp time) per cycle with a target intensity of 20 000 and intensity threshold of 2500, considering charge states 0–5. Active exclusion was used with release after 0.4 min. Advanced collision energy setting with nine points of 11 points of collision energy depending on ion mobility: 0.60 V.s/cm^2 – 15.00 eV; 0.70 V.s/cm^2 – 17.50 eV; 0.80 V.s/cm^2 – 21.00 eV; 0.90 V.s/cm^2 – 26.00 eV; 1.0 V.s/cm^2 – 34.00 eV; 1.10 V.s/cm^2 – 42.50 eV; 1.20 V.s/cm^2 – 52.00 eV; 1.30 V.s/cm^2 – 57.00 eV; 1.40 V.s/cm^2 – 60.00 eV; 1.50 V.s/cm^2 – 62.00 eV; 1.60 V.s/cm^2 – 70.00 eV. Isolation Peak m/z of 2 m/z was used for 700 m/z and 3 m/z was used for 800 m/z. Data were acquired using Timscontrol v6.1.1 and Compass HyStar 6.3.1.8.

Data-independent acquisition proteomic analysis of PIPPI cell lines

For data-independent detection of PIPPI (Supplementary Fig. S2F and G), cell pellets of 30 M cells were lysed in lysis buffer (100 mM HEPES pH 7.6, 150 mM KCl, 1 mM $MgCl_2$) by passing the cell suspension through a 27G needle 10 times using a 1 ml syringe. Subsequently, the cell lysate was incubated on ice for 20 min. The suspension was centrifuged for 4 min at 16 000 × *g* at 4 °C and the supernatant was taken aside. Then the cell pellet was resuspended in lysis buffer and passed through a 27G needle for 10 times. After 10 min incubation on ice, the suspension was centrifuged for 4 min at 16 000 × *g* at 4 °C and the supernatant was mixed with the supernatant obtained from the first centrifugation. Lysate corresponding to 100 µg protein was subjected to digestion. After 5 min denaturation at 95 °C, sodium deoxycholate (DOC) (Sigma–Aldrich) was added to a concentration of 5%. Proteins were reduced for 30 min at 37 °C with 5 mM TCEP (Sigma–Aldrich) and then directly alkylated in the presence of 20 mM iodoacetamide (Sigma–Aldrich) for 40 min at room temperature in the dark. Next, proteins were pre-digested for 4 h at 37 °C with 1 µg Lys-C (Fujifilm Wako Chemicals) and then diluted 1:5 in freshly prepared 0.1 M ammonium bicarbonate buffer to bring the concentration of DOC to 1%. Sequencing-grade Trypsin (Promega) was added at an enzyme: substrate ratio of 1:100 and incubated overnight at 37 °C. The next day, the tryptic digest was stopped, and DOC was precipitated by the addition of FA. To remove DOC precipitates, the digest was centrifuged at 20 000 ×*g* for 10 min and the supernatant was transferred to a fresh tube. This procedure was repeated twice. The cleared digest was loaded onto a 50 mg SepPak C18 column (Waters) which was primed with 100% methanol (Sigma–Aldrich), 80% ACN (Sigma–Aldrich) + 0.1% FA and equilibrated with 1% ACN + 0.1% FA. The flow-through was loaded one more time, and then the column was washed three times with 1% ACN + 0.1% FA. Lastly, peptides were eluted in 35% ACN + 0.1% FA and dried in a SpeedVac for 4 h at 45 °C. For MS measurement peptides were reconstituted in 3% ACN + 0.1% FA.

Peptides were analysed in a data-independent acquisition (DIA) mode with an Exploris 480 (Thermo Scientific) mass spectrometer connected to an EASY-nLC (Thermo scientific) liquid chromatography system operating in nano-flow. Peptides were separated on a 30 cm fused silica column with 75 µm inner diameter packed in-house with 1.9 µm C18 beads (Dr Maisch Reprosil-Pur 120). Peptides were separated along a 2-h nonlinear gradient constituting of a mixture of buffer A (3% ACN + 0.1% FA) and buffer B (90% ACN + 0.1% FA) at a flow rate of 250 nl/min. The DIA method acquired MS1 spectra with a scan range of 350–1650 m/z at a resolution of 120 000 followed by 40 variable MS2 DIA windows with 0.5 m/z overlap at a resolution of 30 000 and a normalized AGC Target of 3000%.

DIA-MS runs were analysed with Spectronaut 16 (Biognosys AG) in direct DIA + search mode. The spectra were searched for Trypsin/P specific peptides with a length between 7 and 52 AA, allowing for two missed cleavages and setting carbamidomethyl(C) as fixed, and oxidation(M) and acetyl(protein, N-term) as variable modifications. Identifications were controlled for a false-discovery rate of 1% on precursor, peptide, and protein level. The Uniprot release from 15.10.2020 of the *Homo sapiens* proteome was used as a reference. To search for peptides corresponding to PIPPI, the fasta sequence of PIPPI was manually added to the reference proteome. The report table in the MSstats format was exported from Spectronaut 16 and further analysed with MSstats (version 4.6.5) in R. Data from Spectronaut was filtered in MSstats using a q-value cutoff of 1% and removing proteins with only one feature. The data were normalized in MSstats using the "equalizeMedians" method and only the top three features were used to build quantities. For differential abundance testing, MSstats fits a linear mixed effects model and applies the Benjamini–Hochberg correction to account for multiple testing. Volcano plots were plotted in R using the ggplot package with significance thresholds set to $\log_2$ fold change ($\log_2$FC) > 1, and adjusted *P*-value <.05. The principal component analysis was done with base R using the MSstats protein quantities as input. The PCA plot was generated with ggplot in R.

DDA proteomic analysis with stress induction (TM, 6-TG)

For proteomic analysis of A375 cells treated with TM and 6-TG (Fig. 3E and F), cells were seeded at a density of 1 million cells per condition and left to adhere overnight. The following day, cells were treated with either vehicle control (DMSO) (Sigma), TM (500 ng/ml), or 6-TG (2 mM) for 24 h. Following treatment, cells were harvested and processed using the EasyPep™ Mini MS Sample Prep Kit (Thermo Scientific) per manufacturer instructions. For MS measurement, peptides were then reconstituted in 0.1% FA.

LC-MS/MS was performed on an Exploris 120 mass spectrometer (Thermo Fisher Scientific), with UltiMate 3000 (Thermo Fisher Scientific) UHPLC system. Separation was performed on PepMap EASYSpray columns at 50 °C. Peptides were eluted in 115 min with a segmented linear gradient of 0%–98% solvent B (solvent A 0% ACN, 0.1% FA; solvent B 80% ACN, 0.1% FA), and a flow rate of 300 nl/min. MS spectra were acquired in DDA mode. MS1 scans were acquired at an Orbitrap Resolution of 120 000 with a Scan Range (m/z) of 380–1500, a maximum injection time of 100 ms and a Normalized AGC Target of 300%. Up to 4 Dependent Scans were taken. For dynamic exclusion, the exclusion duration was set

to 40 s and a mass tolerance of ± 10 ppm. The Isolation Window was set to 1.6 m/z with no offset. A normalized collision energy of 30 was used. MS2 scans were taken at an Orbitrap Resolution of 15 000, with a fixed First Mass (m/z) = 120. Maximum injection time was 22 ms and the normalized AGC Target 50%.

The raw data were analysed using MaxQuant Version 1.6.3.4 [30] and searched against the Uniprot protein database (Human all 2017/11) [31], PIPPI sequences, and the CRAPome database [32]. MaxQuant default settings were used with the two options *Match between runs* and *LFQ intensity reporting* activated. The resulting LFQ intensities were then analysed using the Differential Enrichment analysis of Proteomics data (DEP) package 1.14.0 [33] on RStudio 2022.02.03. After removing common contaminants, imputation was performed using the "MinProb" method without a prior "DEP normalization" of the data. Only proteins present in 3/3 replicates of each pulldown were considered for this analysis.

PIPPI sORFs sequence alignment

The coordinates of the 24 sORFs matching the PIPPI sequence from Supplementary Table S2 were used to retrieve the corresponding nucleotide sequences from Ensembl (release 113 – October 2024), using the script *retrieve_sequences.py*. The retrieved nucleotide sequences were aligned using Benchling's Sequence Alignment Tool (MAFFT v7-based alignment, gap open penalty = 1.53, gap extension penalty = 0). The aligned nucleotide sequences were then translated into protein sequences, compiled in the PIPPI_NPIPfamily_prot.fasta file. Finally, Jalview software (v2.11.4.1) was used to visualize the AA variations, the consensus sequence and the similarity tree. Scripts, files, and sequences are available on the Mendeley Data entry (access details in Data availability section).

Results

Pooled overexpression screens to identify functional sORFs

Although thousands of microproteins have been predicted to exist in mammalian cells, it is highly challenging to define which of these have a function, and to identify their biological role. To efficiently pinpoint functional microproteins, we set out to perform overexpression screens using a large library of sORFs encoding for putative microproteins. We filtered our previously described microprotein database [35], retaining 11 338 sORFs coding for microproteins between 10 and 57 AA (capped by the 200-nucleotide length limit of oligo pool synthesis), and cloned the sORFs into a lentiviral vector (Fig. 1A). With deep sequencing, we confirmed the presence of >10 900 of the sequences in the transduced cells (Supplementary Fig. S1A).

We conducted two parallel screens in human A375 melanoma cells (Fig. 1B): In the first arm, cells were cultured untreated for three weeks to assess whether overexpression of any of the sORFs would promote cellular proliferation, whereas in the second arm of the screen, cells were treated with the cytotoxic compound 6-TG. 6-TG is an analogue of the naturally occurring purine base guanine, and it is mainly used in the clinic for the treatment of acute and chronic myelogenous leukemias. *In vivo,* 6-TG is converted to 6-TG nucleotides, which interfere with a variety of cellular processes involved in nucleic acid synthesis [36]. For example, 6-TG nucleotides are incorporated in the genome of target cells during replication, and repair of the resulting lesions leads to cellular death. In this arm of the screen, cell cultures were continuously treated with 10 μM 6-TG until the surviving cells were able to repopulate the culturing flasks.

For the screen, over 10 million cells were transduced with the pool of sORFs at a low MOI, ensuring that each sORF would be represented at least 1000 times in the cell population. After antibiotic selection to remove uninfected cells, a fraction of cells representing the starting population were collected for later extraction of gDNA. Thereafter, cells were cultured in absence or presence of 6-TG, and, at the end of the screen, gDNA of the different populations was harvested. Finally, sORF cassettes were amplified by PCR and, after library preparation, the abundance of the individual sORFs in the different populations was quantified by Illumina sequencing (Fig. 1B). By plotting the logarithmic fold-change ($\log_2$ FC) in sORFs abundance between the start and the end populations of the cell proliferation screen, we could appreciate that overexpression of the sORFs did not influence cell growth of normally proliferating cells (Fig. 1C). In contrast, with the 6-TG screen, we could see that a large fraction of sORFs were depleted in both replicates (Fig. 1D). In contrast, two sORFs were highly enriched in both replicates upon drug treatment, suggesting that cells expressing these two microproteins are less sensitive toward 6-TG (Fig. 1D). Beside these two hits, other sORFs were enriched in the final population, but in only one of the two replicates (Fig. 1D). This was most likely occurring because of a proliferative advantage that few cells acquired thanks to random mutations, rather than the presence of the sORFs. The top hit in both replicates corresponds to altDDIT3, which is an annotated microprotein encoded by a sORF located upstream of the pro-apoptotic transcription factor CHOP/DDIT3 [37–39]. Previous work on this uORF suggests that it can be translated to a functional microprotein that interacts with the downstream-encoded protein CHOP/DDIT3 [37, 40]. It has also been shown that the uORF encoding for altDDIT3 principally acts *in cis*, either in a peptide-dependent or -independent manner, to downregulate translation of the canonical ORF [41–44]. The second hit encoded a 50 AA microprotein originally suggested based on peptide evidence in a proteomics study [45]. Matching sORF sequences were found in 24 different locations on chr16 and chr18 of the Ensembl 110 human genome annotation, encoded on transcripts of the NPIP gene family (Fig. 1E) [46, 47]. We therefore named this microprotein PIPPI. The NPIP gene cluster, also known as *morpheus* gene cluster or LCR16a core duplicons, encodes many rapidly diversifying NPIP paralogs [48], and is one of the most extreme cases of positive selection observed in primates [46, 47].

To validate the results of the screen, we performed growth competition assays in which the proliferative capacity of a microprotein-overexpressing cell line was directly compared to that of its parental counterpart in a co-culture. In short, a microprotein-expressing cell line was mixed at a 1:1 ratio with a control cell line, and the cell mixture was then either cultured untreated for one week or allowed to recover from a 6-TG treatment for the same period. At the beginning and at the

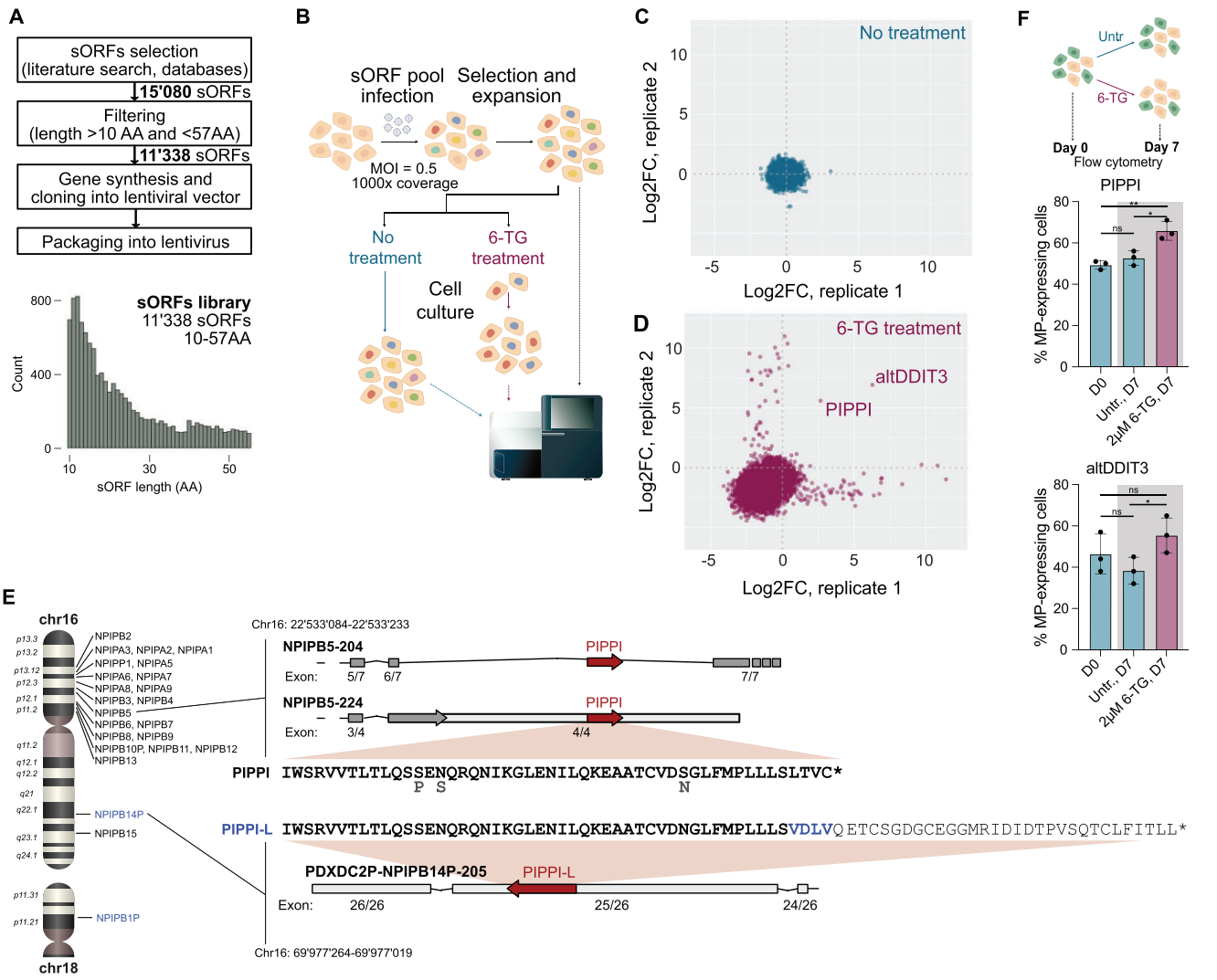


Figure 1. A pooled sORF overexpression screen identifies sORFs conferring resistance to 6-TG treatment. (**A**) Schema depicting the rationale for the design of the sORF library used in this study. (**B**) Workflow of the pooled overexpression screens performed to identify sORFs modulating cell proliferation or resistance to 6-TG. (**C, D**) Scatterplots depicting the results of the cell proliferation screen (**C**) and the 6-TG screen (**D**). The $\log_2$FC in abundance of each sORF between the start and the end of the screen is plotted for the two biological replicates. (**E**) Chromosomal location of *nuclear pore complex interacting protein* (NPIP) genes encoding for the PIPPI sORF (left), and exemplary alternative splicing transcript variants annotated in Ensembl. PIPPI sORFs (PIPPI, PIPPI-L) are shown in red arrows, and the consensus sequence is shown in bold. Minor sequence variants are depicted in grey below the PIPPI consensus sequence. See also Supplementary Fig. S1C for sequence alignment and Supplementary Fig. S1D for example RNA-seq coverage. (**F**) Top: Schema depicting the experimental setup for the growth competition assays. Bottom: Bar charts representing the results of the growth competition assays performed to validate the results of the 6-TG screen. Where indicated (violet bars), cells were treated for 24 h with 2 µM 6-TG. Height of the bars represents the fraction of microprotein-expressing cells present in the total cell population. Values were averaged based on three independent biological replicates (black circles). *P*-values were calculated using unpaired Student's *t*-test (ns = $P > .05$; * = $P < .05$; ** = $P < .01$).

end of the assay, the ratio between GFP-positive and -negative cells was assessed by flow cytometry (Fig. 1F). Validating the screening results, expression of untagged PIPPI did not provide a proliferation advantage over GFP-only-expressing cells in untreated conditions, but cells expressing the microprotein proliferated better than the parental cell line after treatment with 2 µM 6-TG (Fig. 1F). Similarly, cells expressing altDDIT3 proliferated better than parental cells after a 6-TG treatment but not in untreated conditions (Fig. 1F). The same observation could be made when performing the assay using a cell line expressing a GFP-tagged version of PIPPI and treating them with a higher dose of 6-TG (10 µM 6-TG, the same dose used for the screen). Interestingly, when the same strategy was applied to altDDIT3, we could observe that including the GFP tag inhibited the protective activity of altDDIT3 (Supplementary Fig. S1B). Altogether, our data support previous reports indicating that altDDIT3 can be functional as a microprotein and suggests that it could have a role in inhibiting the pro-apoptotic effects of CHOP/DDIT3, while PIPPI appeared to exert a similar effect through a yet unknown mechanism. We therefore chose to focus on PIPPI for subsequent characterization.

PIPPI is a novel microprotein localizing to the ER

The coding sequence for PIPPI originates from a highly duplicated region of the human genome. 23 sORFs matching the PIPPI microprotein sequence were found on chr16 (Fig. 1E,

Supplementary Fig. S1C, and Supplementary Information) [49], all encoded by NPIP genes/pseudogenes of the 20 kbp core duplicon LCR16a [46]. One additional match mapped to the pseudogene/long-noncoding RNA (lncRNA) NPIPB1P on chr18. Further, 11 matches were found on alternative chromosome scaffolds. Members of the NPIP gene family can be subdivided in two subfamilies, NPIPA and NPIPB, encoding for structurally different proteins, with the NPIPB subfamily being exclusively expanded in chimpanzee, human, and gorilla [47, 50]. In the majority of instances, PIPPI sORF is located in the last intron of the reference transcript, but in certain splice variants, the PIPPI sORF is included in the mature transcripts (Fig. 1E). NPIP family transcripts appear to exhibit diverse tissue-specific expression patterns [48]. Inspecting published poly-A RNA-seq from A375 cells [35], we confirmed that the PIPPI sORF-containing introns show read coverage and are hence included in many NPIP transcripts variants (Supplementary Fig. S1D). The AA sequence of all 24 PIPPI paralogs is largely identical, with two minor single AA variants S13P and N15S in one and seven paralogs, respectively (Fig. 1E and Supplementary Fig S1C). In two pseudogenes/lncRNAs, PDXDC2P-NPIPB14P and NPIPB1P, the NPIP ORF is lost and a longer isoform of PIPPI, named here as PIPPI-L, is encoded in the last exon (Fig. 1E). When we exogenously expressed the two GFP-tagged microproteins in A375 cells, we noticed that PIPPI was more stable than PIPPI-L (Supplementary Fig. S2A), we therefore decided to focus on this isoform for further characterization of the protein.

To characterize the cellular role of PIPPI, we overexpressed tagged versions of the microprotein in various cell lines, confirming its expression and subcellular localization in both stable A375, RPE-1 and U2OS cells (Fig. 2A and Supplementary Fig. S2A and C), as well as transiently transfected COS-7 and HEK293T cells (Fig. 2A and Supplementary Fig. S2D). We decided to work with different tags to ensure that these tags, despite their sizes and biophysical properties, would not affect the subcellular localization and function of the microprotein. When a single HA tag was included at the C-terminus of the microprotein, PIPPI-HA, the construct appeared to be particularly lowly expressed and/or unstable, since we were not able to detect the tagged microprotein in transiently transfected cells by immunoblotting (Supplementary Fig. S2D). However, fused to a STELLA tag [51] leader sequence, we were able to detect the STELLA-PIPPI-HA microprotein via bioorthogonal label and HA-tag (Supplementary Fig. S2D). We observed STELLA-PIPPI-HA enrichment in the endoplasmic reticulum (ER) when expressed in COS-7 cells, as highlighted by colocalization with the lectin Con A, a well-established marker of this organelle (Fig. 2A). A stably expressed PIPPI-GFP fusion also colocalized with the ER marker in A375 (Fig. 2A), RPE-1 and U2OS cells (Supplementary Fig. S2B and C). To ascertain whether PIPPI was associated to the cytoplasmic side of the ER membrane or was present within the ER lumen, we engineered a glycosylation site in PIPPI (PIPPI^{Q12N}-GFP). When transiently expressed in HEK293T cells, this construct displayed an additional PIPPI band that migrated slower in SDS–PAGE, and could be reverted by treatment with the de-glycosylase enzyme PNGase F (Fig. 2B). Altogether, these results suggest that a fraction of PIPPI is localizing within the glycosylation-supporting environment of the ER lumen [52], despite the fact that the microprotein does not possess a canonical ER localization sequence.

Validation of endogenous PIPPI expression

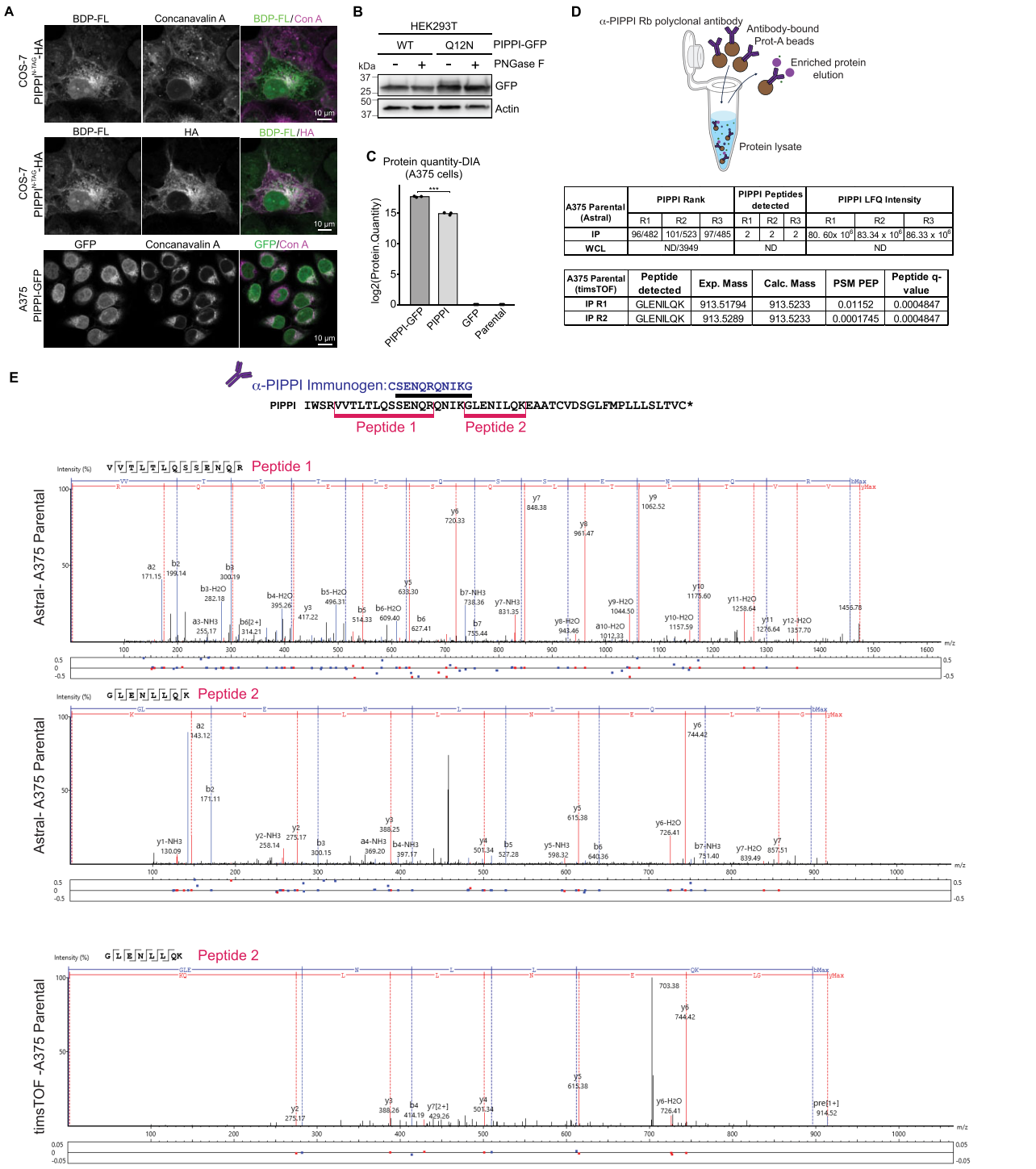
To gain insight on the endogenous expression, localization, and regulation of PIPPI, we raised a rabbit antiserum against the epitope sequence SENQRQNIKG, chosen based on uniqueness in the human genome and predicted antigenicity. The affinity-purified antibody robustly detected exogenously expressed PIPPI-GFP (Supplementary Fig. S2A). However, we detected neither endogenous PIPPI in parental A375 cells, nor overexpressed (tagless) PIPPI. Instead, the antibody showed various unspecific bands in the western blot (Supplementary Fig. S2A). This could be ascribed either to an inefficient transfer, or high threshold for detecting untagged PIPPI, since large amounts of the synthetic PIPPI peptide were needed to be spiked into the lysate to gain a western blot signal (Supplementary Fig. S2E).

In a parallel effort, we sought to validate the presence of endogenous PIPPI by MS. We performed DIA deep proteomics on whole cell lysates of parental A375, PIPPI, and PIPPI-GFP overexpressing cells (Fig. 2C) and found robust evidence for overexpressed PIPPI and PIPPI-GFP based on two independent peptides (see Fig. 2E), but no PIPPI peptides were detected in the parental A375 lysate. Surveying the proteome-wide differences in parental A375, PIPPI, PIPPI-GFP, or GFP-expressing cells in the same DIA data, we found that any effects induced by exogenously expressing PIPPI were overall mild, with very few other proteins being significantly increased or decreased (Supplementary Fig. S2F and G andSupplementary Table S3).

We hypothesized that, despite its limited performance on a western blot, our polyclonal antibody could be used to enrich for endogenous PIPPI as a pre-fractionation step for a proteomics experiment (Fig. 2D). We performed PIPPI immunoprecipitation (IP) from A375 cells on Protein A-magnetic beads followed by LC-MS/MS with DDA acquisition on two high-end mass spectrometers, Thermo Orbitrap Astral and Bruker timsTOF. In addition to the parental A375 cell line, we also included A375 PIPPI and A375 PIPPI GFP cell lines in this experimental set up. We were able to detect two tryptic peptides, deriving from the PIPPI consensus sequence in wildtype A375 cells (Fig. 2D and E). While undetectable in whole cell lysates, PIPPI ranked amongst the top ~20% most abundant proteins after the antibody-based enrichment. On the timsTOF system, we were able to detect one of the peptides with high confidence (Fig. 2D and E). The same peptides were also identified in pre-enriched lysates from PIPPI and PIPPI-GFP overexpressing cells (Supplementary Fig. S2H). In summary, antibody-based pre-enrichment provided us with a means to specifically enrich the cellular proteome for our microprotein of interest, a strategy that could be generally applicable for enhancing detection of endogenous microprotein species.

The microprotein PIPPI enables cells to overcome ER stress

Based on the results of our initial screen, we hypothesized that PIPPI could be involved in the repair of nuclear DNA lesions. However it is also known that 6-TG treatment can result in mitochondrial DNA damage, mitochondrial dysfunction, and the activation of the unfolded protein response (UPR), a stress-response pathway that cells activate to relieve the accumulation of misfolded proteins in the ER [53–56]. While we did not find evidence that PIPPI alleviated cellular stress in response



A375 Parental (Astral)	PIPPI Rank			PIPPI Peptides detected			PIPPI LFQ Intensity		
	R1	R2	R3	R1	R2	R3	R1	R2	R3
IP	96/482	101/523	97/485	2	2	2	80. 60x 10^6	83.34 x 10^6	86.33 x 10^6
WCL	ND/3949			ND			ND		

A375 Parental (timsTOF)	Peptide detected	Exp. Mass	Calc. Mass	PSM PEP	Peptide q-value
IP R1	GLENILQK	913.51794	913.5233	0.01152	0.0004847
IP R2	GLENILQK	913.5289	913.5233	0.0001745	0.0004847

Figure 2. PIPPI is a novel microprotein localizing to the ER. (**A**) STELLA tag was used to define the subcellular localization of PIPPI-HA in COS-7 cells. The BODIPY-FL-tetrazine dye (BDP-FL) was reacted with TCO∗K incorporated in the PIPPI-HA sequence. The ER is stained using Con A. Images presented in Fig. 2A are overlapped to highlight the extent of colocalization between the BDP-FL and HA signals, as well as the BFP-FL and Con A signals. In the lowest panel are representative confocal images of A375 cells stably expressing a PIPPI-GFP transgene. Con A staining was included to visualize the ER. (**B**) A PIPPI-GFP point mutant engineered to allow N-glycosylation were transiently expressed in HEK293T cells. After lysis, samples were treated with the recombinant amidase PNGase F and analysed by immunoblotting. Actin was used as the loading control. (**C**) MS (DIA) was used to quantify the abundance of PIPPI and PIPPI-GFP in whole cell lysates from A375 cells, represented by the bar plot. *P*-values were calculated using unpaired Student's *t*-test (ns = $P > .05$; ∗ = $P < .05$, ∗∗ = $P < .01$. ∗∗∗ = $P < .001$). (**D**) Schema showing the pre-enrichment strategy, using PIPPI antibody-bound Protein-A magnetic beads. The tables below summarize the identification of PIPPI tryptic peptides in Astral (top table, three replicates) as well as timsTOF (bottom table). (**E**) Schema of tryptic peptides identified by MS (Peptide 1, Peptide 2) as well as the immunogen used to raise the custom-made PIPPI antibody used for pre-enrichment. MS2 spectra from A375 parental cells in both Astral and timsTOF are presented.

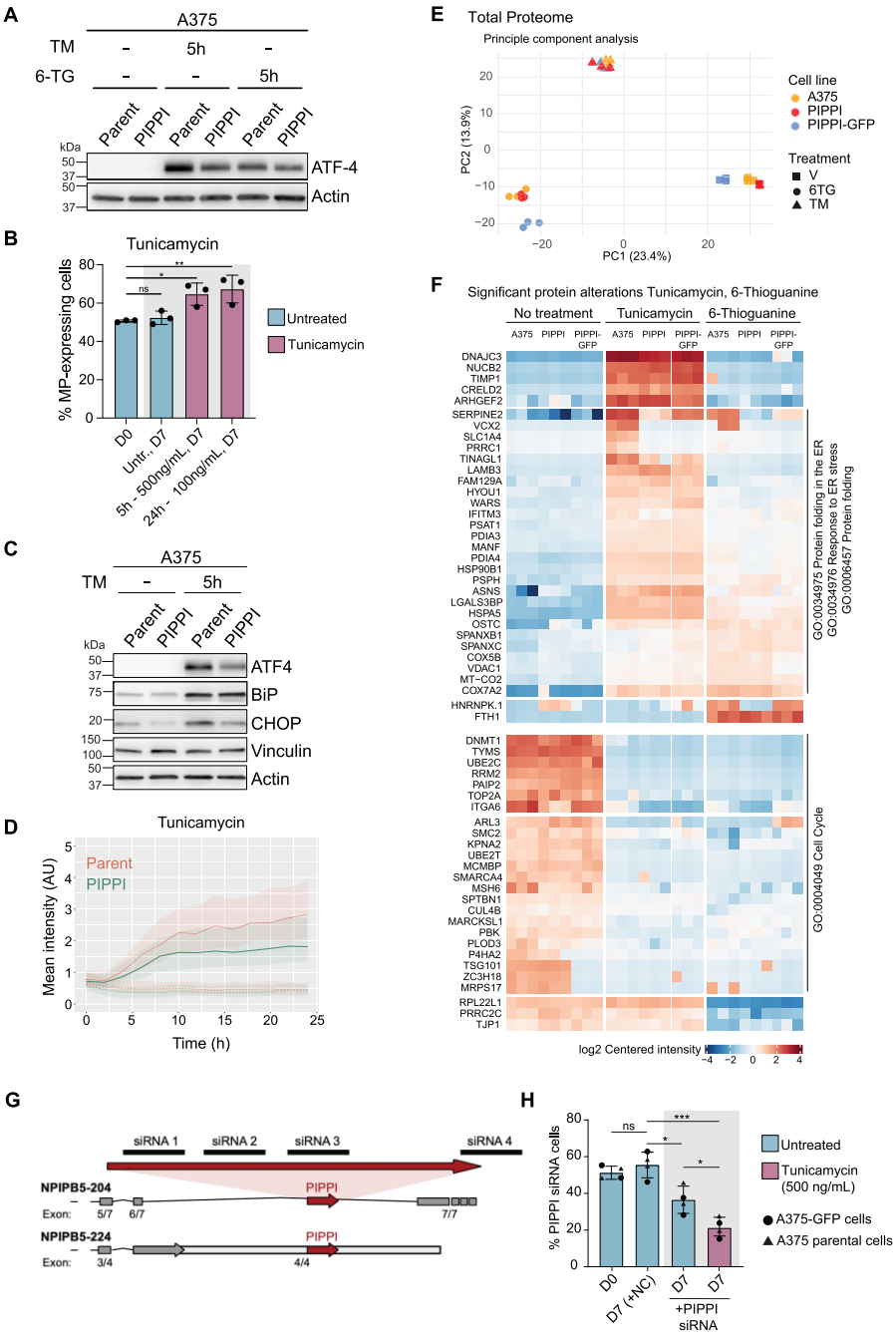


Figure 3. PIPPI is involved in the ER stress response. (**A**) Parental and PIPPI-expressing A375 cells were treated for 5h with either 500 ng/ml TM or 5 μM 6-TG. The level of ATF4 induction was assessed by immunoblotting. Actin was used as a loading control. (**B**) Bar chart representing the results of the growth competition assays performed to assess the effects of TM on cellular proliferation of parental versus microprotein-expressing cells. Height of the bars represents the fraction of PIPPI-expressing cells present in the total cell population. Values were averaged based on three independent biological replicates (circles). *P*-values were calculated using unpaired Student's *t*-test (ns = $P > .05$; ∗ = $P < .05$, ∗∗ = $P < .01$). (**C**) Parental and PIPPI-expressing A375 cells were treated with 500 ng/ml TM for 5 h. After lysis, the expression of different proteins involved in ER stress response were assessed by immunoblotting. Actin was used as a loading control. (**D**) Parental and PIPPI-expressing A375 cells, engineered to express a fluorescent ATF4 reporter, were treated with 500 ng/ml TM, and imaged every 30 min over the course of 24 h. Wells with untreated cells were included and are reported in the graph as dashed lines. Plotted are the median, first quartile and third quartile values that were calculated using the mean intensity of at least 300 cells. (**E**) Principal component analysis of whole cell lysate MS performed on A375 parental, PIPPI, and PIPPI-GFP expressing cells treated with either vehicle (DMSO), TM (500 ng/ml, 24 h), or 6-TG (2 mM, 24 h). (**F**) Clustered heatmap representing $\log_2$ centered intensity of significantly altered proteins comparing 6-TG and TM treatments with vehicle control, with gene ontology analysis of the proteins commonly deregulated in 6-TG and TM treatments. (**G**) Schematic showing the siRNA target locations used for knocking down the PIPPI sORF region in A375 cells. (**H**) Bar graph representing competitive growth between A375 parental (GFP−) and A375-GFP (GFP+) cells transfected with siRNAs targeting PIPPI or a nontargeting NC, and further treated with either vehicle (DMSO) or TM (500 ng/ml). Shown is proportion of GFP− cells transfected with PIPPI siRNAs (triangles) competing with GFP + cells transfected with NC or GFP + cells transfected with PIPPI siRNAs (circles) competing with GFP− cells transfected with NC. Statistical significance was assessed by two-way ANOVA followed by Tukey's multiple comparisons test. Adjusted *P*-values are indicated as follows: (ns = $P > .05$; ∗ = $P < .05$, ∗∗ = $P < .01$, ∗∗∗ = $P < .001$). Values were averaged based on four independent biological replicates (2 GFP+, and 2 GFP− replicates, to account for GFP labelling bias).

to DNA-damaging agents camptothecin, etoposide, and cisplatin (data not shown), we ascertained that PIPPI modulated the UPR: When treating parental and PIPPI-expressing A375 cells for 5 h with either 6-TG or TM, a compound inhibiting the first step of N-linked glycosylation [57], we confirmed that both drugs triggered a stabilization of activating transcription factor 4 (ATF4), a central player of the UPR. Interestingly, in these experiments, we also noticed that ATF4 activation in PIPPI-expressing cells was attenuated compared to parental A375 (Fig. 3A). We further performed growth competition assays, in which cells were treated with a sub-lethal dose of TM. Confirming the involvement of PIPPI in an ER stress response pathway, PIPPI-expressing cells were able to grow better than their parental counterparts after either a short (5 h) or long (24 h) treatment with TM (Fig. 3B), with a similar outcome in PIPPI-GFP expressing cells (Supplementary Fig. S3A).

To better understand the effects of PIPPI on cancer cells' ability to deal with misfolded proteins, we assessed the induction of some key components of the UPR. The chaperone BiP/HSPA5 plays a critical role in sensing the accumulation of misfolded proteins in the ER and, upon binding to unfolded proteins, further activating transcription factors ATF6 and IRE1 [58]. ATF6 and IRE1 promote the upregulation of ER chaperones, such as BiP itself, and folding enzymes, such as protein disulfide isomerase (PDI). BiP also activates the PKR-like ER kinase PERK, one of the four kinases constituting the integrated stress response (ISR), which in turn phosphorylates the eukaryotic initiation factor 2 alpha (eIF2α) [59]. As a result of eIF2α phosphorylation, the rate of translation initiation is decreased to reduce general protein synthesis and thereby alleviate the situation of stress. Under these conditions, a subset of mRNAs, which include the one encoding for ATF4, are selectively translated to promote recovery and survival [60]. Finally, if the stress situation is not promptly resolved, prolonged activation of the UPR results in the upregulation of the C/EBP Homologous Protein (CHOP/DDIT3) transcription factor, which initiates apoptotic cell death [39]. Interestingly, despite both cell lines displaying a similar upregulation of BiP which is indicative of the cells being subjected to similar levels of ER stress, PIPPI-expressing cells had a dampened downstream signalling, resulting in a weaker stabilization of both transcription factors ATF4 and CHOP (Fig. 3C). Importantly, the same phenotype could also be recapitulated in A375 cells expressing either PIPPI-GFP or PIPPI-L (Supplementary Fig. S3B and C). To confirm this observation using an independent approach, we engineered A375 cells to express a fluorescent reporter for ATF4 induction [61], and imaged the living cells for 24 h after treatment with TM. Two hours after TM treatment, the levels of ATF4 started increasing in both parental and PIPPI-expressing cells, but from the beginning the PIPPI curve was less steep than the parental one. Moreover, ATF4 levels plateaued earlier in PIPPI cells, corroborating the initial observation that these cells have a dampened activation of the UPR after exposure to ER stress (Fig. 3D). Further, experiments performed in human U2OS osteosarcoma cells confirmed that the dampened activation of the UPR is not an A375-specific phenotype, but can be recapitulated in cells of different origin (Supplementary Fig. S3D and E).

To further understand the relationship between TM- and 6-TG-induced stress and the role of PIPPI, we performed a proteomics survey of parental, PIPPI- and PIPPI-GFP A375 cells treated with TM or 6-TG. Notably, the majority of significantly up- and downregulated proteins in the response to TM and 6-TG were shared (Fig. 3E and F and Supplementary Table S4); common upregulated proteins were enriched in GO terms 'protein folding' and 'ER stress response', whereas downregulated proteins were indicative of a cell cycle arrest (Fig. 3F and Supplementary Fig. S3F and G). The overall proteome response was highly similar between parental, PIPPI, and PIPPI-GFP overexpressing cells (Fig. 3E), albeit many of the key ER-stress induced factors, including ATF4, BiP, CHOP, were too low abundant to be detected in any of the treatments in our proteomics dataset of ~2000 quantified proteins. A known group of genes downstream of ER stress are AA transporters [62, 63], and intriguingly we found one of the previously described ATF4 targets, SLC1A4, to be highly upregulated in parental A375, but not PIPPI and PIPPI-GFP expressing cells upon TM treatment (Fig. 3F). This may indicate that, through attenuation of ATF4 activity, PIPPI also attenuates activation of some of the downstream stress response genes.

Given the collective evidence from the overexpression studies above, pointing to a role of PIPPI in dampening the stress response, we hypothesized that depleting endogenous PIPPI would in turn sensitize cells to ER stress. Given the complexity of the NPIP gene cluster, various challenges existed to perform a loss-of-function experiment: it is unclear which NPIP paralog expresses a transcript retaining the intronic PIPPI sORF, and whether PIPPI is produced from a single or multiple NPIP genes, or the pseudogenes encoding for PIPPI-L (Fig. 1E). The high sequence similarity between paralogs made it impossible to target specific NPIP transcripts e.g. with CRISPR/Cas9, and any knockout or knockdown of the main NPIP transcript would invariably affect expression of the NPIP proteins with unknown consequences. However, we reasoned that targeting RNA interference (RNAi) to the PIPPI sORF would enable depletion of PIPPI while sparing the main NPIP transcript for which the targeted region lies in an intron. The PIPPI sORF and surrounding sequences are virtually identical amongst all occurrences, allowing us to design four short interfering RNAs (siRNAs) spanning all PIPPI sORF (Fig. 3G). None of the four siRNAs had significant matches elsewhere in the human genome (Supplementary Fig. S3H). We performed co-culture experiments mixing parental and GFP-expressing A375 cells after transfecting each with a pool of the PIPPI siRNAs or nontargeting NC siRNAs, respectively (Fig. 3H). In 2/4 replicates, the labeling of PIPPI and NC siRNA-transfected cells was swapped (see the 'Materials and methods' section for co-culture experiment details). By tracking the fraction of GFP− and GFP + cells over time, with or without TM treatment, we were able to determine if PIPPI siRNA affected cell growth and stress sensitivity compared to the control siRNA. Strikingly, PIPPI knockdown cells showed a significant growth defect compared to NC-transfected cells, which was further exacerbated after TM treatment (Fig. 3H). These results support a protective role for endogenous PIPPI in the cellular response to ER stress, although it cannot be ruled out that the knockdown of PIPPI-containing transcripts has wider effects on NPIP main ORFs or noncoding functions of the affected NPIP genes/pseudogene transcripts. In summary, PIPPI-expression correlates with a dampened response when experiencing ER stress and, consequently, provides cells with a better chance to survive and proliferate under stress conditions.

Dissecting PIPPI's interactome

The presence of PIPPI at the ER, including the ER lumen (Fig. 2A and B), in combination with its apparent ER-related function (Fig. 3), led us to hypothesize that PIPPI may be exerting its function as stress mediator in the ER. To identify cellular interaction partners of PIPPI, we performed co-immunoprecipitation (co-IP) experiments followed by LC-MS/MS with the PIPPI-GFP fusion and controls in A375 and U2OS cells. Our rationale for using a strong affinity tag was the limited specific enrichment with the PIPPI antibody (Fig. 2D). Capturing PIPPI-GFP with GFP-nanobody-coupled beads, we identified several ER-resident (BiP/HSPA5, P4HB, DNAJC10, ERLIN1) or ER-golgi trafficking proteins (ERp44, TMED10) that were enriched in the PIPPI-GFP sample in both A375 and U2OS cells (Fig. 4A–C and Supplementary Table S5). STRING network analysis of the top 10 putative PIPPI interactors highlighted proteins belonging to biological processes significantly enriched in the network. For example, blue nodes indicate proteins involved in the "response to ER stress" (FDR = 0.00054) and green nodes correspond to proteins related to "protein folding in the ER" (FDR = 0.00054). In addition to this, nodes associated with cellular components significantly enriched in the network, in this case the "ER-Golgi intermediate compartment" (ERGIC), are highlighted in red (FDR = $7.85e^{-6}$) (Fig. 4D).

We further validated the interaction of PIPPI-GFP with BiP and the Endoplasmic Reticulum Protein 44 (ERp44) in U2OS cells (Fig. 4E and F). BiP is a chaperone involved in the folding and assembly of proteins in the ER, and its main function is to recognize and bind unfolded proteins in the ER. Hence, we were not able to distinguish if PIPPI-GFP is a folding client of BiP or binds BiP in another way. BiP protein levels increase when unfolded proteins accumulate in the ER, but we did not observe such effect by PIPPI overexpression, corroborating that PIPPI itself does not induce or increase protein folding stress in the ER (e.g. Fig. 3C and Supplementary Fig. S3C). ERp44, the other top interactor, is a member of the PDI family of ER proteins, which principally functions as a pH- and Zinc-dependent chaperone along the secretory pathway. In fact, ERp44 cycles between the ER and the Golgi, controlling the secretion of disulfide-linked oligomeric proteins and ensuring the correct localization of ER enzymes, lacking localization signals [64]. Besides these canonical roles, ERp44 has been shown to interact and regulate a brain-specific subtype of the Inositol 1,4,5-trisphosphate receptors (IP_3Rs), which are intracellular channels controlling calcium release from the ER [65].

To better understand the relationship between PIPPI and its interaction partners, we generated a panel of cell lines expressing PIPPI-GFP truncations (Fig. 4G) and analysed them by co-IP and confocal microscopy (Fig. 4H and I). The most striking phenotype was observed for $PIPPI^{\Delta 44–51}$, lacking its C-terminus, resulting in the loss of almost all interaction partners (Fig. 4H), and a pan-cellular localization (Fig. 4I). A diffuse localization was also observed for the N-terminal mutant $PIPPI^{\Delta 2–9}$, despite which this mutant seemed to retain some of the interaction partners (Fig. 4H). Interestingly, both of these PIPPI mutants lost their interaction with TMED10. Thus, we hypothesize that in normal conditions, the Golgi complex-localizing protein channel TMED10 is shuttling PIPPI into the lumen of the secretory pathway, where it is retained principally thanks to the action of ERp44. Another interesting phenotype could be observed for $PIPPI^{\Delta 28–35}$, which appeared to be particularly unstable (Fig. 4H and I). Finally, the Δ10–17, Δ20–27, and Δ36–43 mutants appeared to retain their capability of localizing in the ER and Golgi complex, but their distribution within the compartments could be affected by their differential ability to interact with ERp44, BiP, and ERdj5. Interestingly, mutants Δ20–27 and Δ36–43 seemed to be able to reach the ER, as judged by their capability to interact with BiP and ERdj5, despite not interacting with ERp44 (Fig. 4H). This indicates that there is another factor able to relocate PIPPI from the Golgi to the ER. Further, mutants of the C-terminal domain of PIPPI, where the only two cysteines of PIPPI are located, interact less promptly with ERdj5. We confirmed the interaction of PIPPI-GFP with TMED10 and ERdj5/DNAJC10, also showing distinct dependencies on interacting regions within PIPPI (Fig. 4H). TMED10 is a type I membrane protein which localizes to the plasma membrane and Golgi and is involved in protein trafficking [66]. More recently, it has also been shown that TMED10 can mediate the uptake of leaderless secretory proteins into the ERGIC, favouring their secretion via an unconventional protein secretion pathway [67]. ERdj5 is an ER disulfide reductase capable of both promoting correct folding of proteins by removing nonnative disulfide bonds, but also for initiating the degradation of misfolded proteins [68, 69]. To summarize, we showed that PIPPI interacts with different proteins within the secretory compartment, predominantly ERp44 and BiP. In addition, we tentatively delineate the role of each interactor for PIPPI localization and function.

The ability to interact with ERp44 is a prerequisite for PIPPI's function

To further validate the interaction between PIPPI and ERp44, we performed reverse co-IP by immunoprecipitating HA-ERp44 in lysates obtained from PIPPI-GFP A375 cells engineered to co-express HA-ERp44 (Supplementary Fig. S4A). We performed a similar co-IP experiment after co-overexpressing ERp44 and PIPPI-GFP in HEK293T cells (Fig. 5A). In this experiment, we included the $ERp44^{C29V}$ mutant which is unable to interact with its client proteins [70]. The mutated cysteine C29, which is part of the N-terminal thioredoxin-like domain of ERp44 (CRFS), is in fact necessary to form mixed disulfide bonds with client proteins [70]. Interestingly, the C29V mutant of ERp44 was not abrogating the interaction with PIPPI, suggesting that PIPPI is not a canonical client protein of ERp44 (Fig. 5A). To clarify the importance of (mixed) disulfide bonds for PIPPI function, and encouraged by the observation that $PIPPI^{\Delta 28–35}$ and $PIPPI^{\Delta 44–51}$ severely affect the microprotein's stability and localization, we decided to point-mutate to alanine the two cysteines present in PIPPI (C35A and/or C51A). The double mutant $PIPPI^{C35A–C51A}$ appeared to be particularly unstable, but the single point mutants could be expressed in cells and were slightly more stable than the *wild type* microprotein, suggesting that the two cysteines forming an intramolecular disulfide bond is not a prerequisite for proper folding and stabilization of the protein. When we tested how the cysteine mutants affected the interactome of PIPPI, we noticed that both point mutants completely abrogated the interaction with ERp44 and increased the fraction of the protein that remains bound to TMED10 (Fig. 5B). Similar to what was observed with the truncation mutants, inhibiting the ability of PIPPI to interact with ERp44 decreased the amount of the microprotein that reaches the ER,

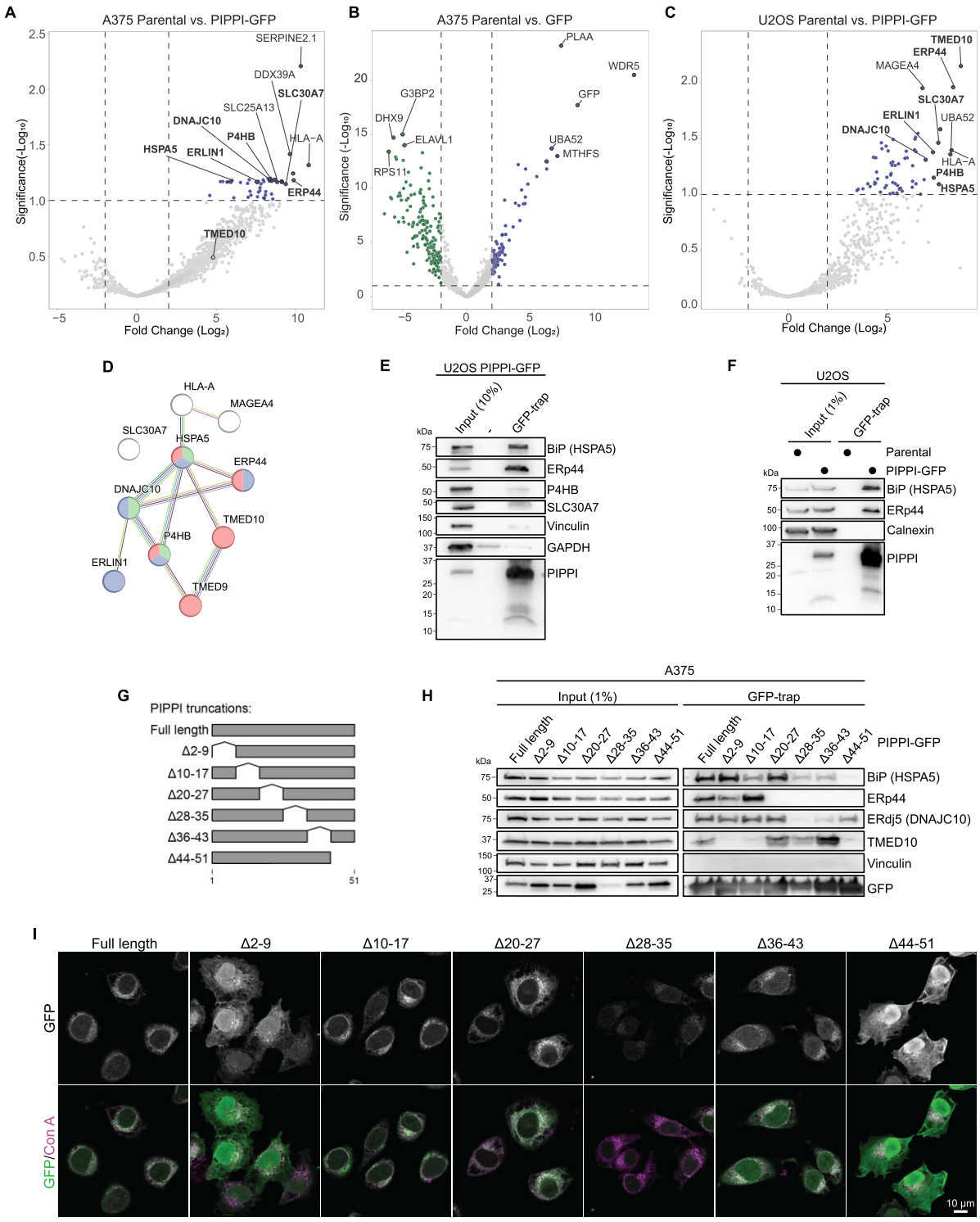


Figure 4. Identification of PIPPI interaction partners. (**A, B**) Volcano plots of co-immunoprecipitation LC-MS/MS (co-IP/MS) experiment performed with parental, GFP, and PIPPI-GFP-expressing A375 cells. The experiment was conducted in two biological replicates, each analysed with three technical replicates. Thresholds are set at $\log_2$FC = 2 and DEP adjusted *P*-value = 0.1 (dashed lines). Proteins enriched in PIPPI-GFP lysates are clustering in the right half of the plot, top enriched proteins (assessed by Manhattan distance) are labelled in the plot. (**C**) Volcano plot of co-IP/MS experiment performed with parental, and PIPPI-GFP-expressing U2OS cells. Commonly enriched proteins are highlighted in bold in A375 and U2OS plots. (**D**) STRING network of the putative top 10 PIPPI interactors identified by co-IP/MS in U2OS cells. Nodes belonging to biological processes significantly enriched in the network are highlighted in blue (Response to ER stress, FDR = 0.00054) and in green (Protein folding in the ER, FDR = 0.00054). Nodes associated to cellular components significantly enriched in the network are highlighted in red (ERGIC, FDR = $7.85e^{-6}$). (**E**) U2OS-PIPPI-GFP lysates were immunoprecipitated using a GFP-trap and analysed by immunoblotting to validate the co-IP/MS results. (**F**) Lysates obtained from parental and PIPPI-GFP-expressing U2OS cells were immunoprecipitated using a GFP-trap and analysed by immunoblotting. (**G**) Schema depicting the different PIPPI-GFP truncations that were tested in Fig. 4H and I. (**H**) Lysates from A375 cells stably expressing different deletion mutants of PIPPI were immunoprecipitated using GFP-trap, and analysed by immunoblotting to identify regions in PIPPI responsible for protein–protein interactions. (**I**) Confocal images of A375 cells stably expressing a panel of PIPPI-GFP truncations. Con A staining was included to visualize the ER.

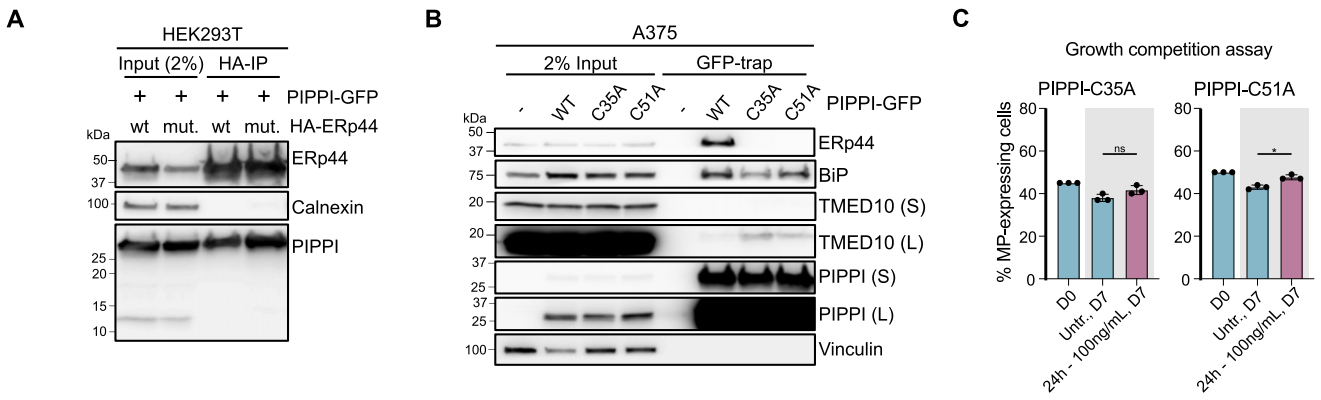


Figure 5. PIPPI exerts its protective function through the PDI ERp44. (**A**) PIPPI-GFP was transiently co-overexpressed with either wild-type HA-ERp44 or a *CRSF mutant* of HA-ERp44 in HEK293T cells. Lysates were subjected to immunoprecipitation using HA-beads and analysed by immunoblotting. (**B**) Lysates obtained from parental, *wild type* PIPPI-GFP-, *C35A* PIPPI-GFP-, and *C51A* PIPPI-GFP-expressing A375 cells were subjected to immunoprecipitation using GFP-trap and analysed by immunoblotting. (S) indicates a short exposure, whereas (L) indicates a long exposure of the same antibody. (**C**) Bar chart presenting the results of the growth competition assays performed to assess the effects of TM on the proliferation of *C35A* PIPPI-GFP- and *C51A* PIPPI-GFP-expressing A375 cells. Height of the bars represents the fraction of PIPPI-expressing cells present in the total cell population. Values were averaged based on three independent biological replicates (circles). At least 10 000 cells were measured for each sample. *P*-values were calculated using unpaired Student's *t*-test (ns = $P > .05$; $* = P < .05$).

as implied by the lower amount of PIPPI binding to BiP (Fig. 5B). Finally, we investigated how these two PIPPI mutants influenced cell proliferation upon treatment with TM, and we observed that overexpression of neither of them was helping cells coping with ER stress to the extent measured with the *wild type* PIPPI microprotein (Fig. 5C). In line with this, the dampened activation of the UPR observed in PIPPI cells was less accentuated in cells expressing the one or the other cysteine mutant, suggesting that interaction with ERp44 is critical for PIPPI functionality (Supplementary Fig. S4B).

Discussion

In this study, we presented a pooled cDNA overexpression screen, capable of identifying microproteins with specific cellular phenotypes, out of an sORF library of over 10 000 candidates. One disadvantage of a pooled screen is the inability to validate how many of the supplied sORFs produce stable microprotein products. Since the majority of our sORF candidates were derived from ribosome profiling evidence, it may be expected that many of these represent sORFs for translation control that do not produce a stable functional microprotein. However, the pooled format does not suffer from carrying along many nonfunctional clones, hence providing the bandwidth to include large numbers of candidates in order to "find the needle in a haystack". In a proof-of-concept screen performed to uncover novel factors promoting cell survival upon exposure to 6-TG, we identified the known microprotein altDDIT3 and a novel, uncharacterized microprotein, which we named PIPPI. PIPPI is encoded within a highly duplicated region of the human genome, the NPIP gene family [46, 71]. In all NPIP genes, the PIPPI sORF is found downstream of the main NPIP ORF, and located in an intron of the most annotated NPIP transcripts. Nevertheless, poly-A RNA-seq data suggests that the PIPPI sORF is part of splice variants that include the intron.

While PIPPI was first suggested on the basis of proteomics evidence in K562 cell lysates [45], we were not able to detect PIPPI by DDA (data-dependent acquisition) shotgun proteomics in a range of cell lysates, including A375, K562, U2OS, HEK293T, even when applying small protein fractionation methods (data not shown). Raising a polyclonal antiserum against a predicted immunogenic sequence within PIPPI did not yield highly specific antibodies that could be used for general western blot, immunofluorescence, or immunohistochemistry applications, but provided us a means to enrich PIPPI from cell lysates, and ultimately allowed us to detect tryptic peptides from endogenously produced PIPPI.

Validating the role of PIPPI in 6-TG resistance, we uncovered functional and biochemical links between PIPPI and the unfolded protein response (UPR). Co-culture experiments provided us a sensitive assay for assessing cell fitness under normal growth and ER stress conditions. Overexpressing PIPPI or PIPPI-GFP did not affect normal growth in comparison to the parental A375 cells, but enhanced growth after TM and 6-TG treatments. In turn, RNAi targeting the PIPPI sORFs increased sensitivity to TM treatment, suggesting a positive correlation between PIPPI levels and resistance to ER stress. Knockdown of PIPPI sORFs reduced cell proliferation already under normal growth conditions, an interesting finding given that the NPIP proteins have not been assigned an essential function despite the strong positive selection observed for the gene cluster [48, 50]. A limitation of our study is that we could not delineate which (one or more) instances of the sORF could produce the microprotein, and how its transcription and translation is regulated, together or separately from the NPIP gene products. Hence, further studies will be needed to dissect the functional elements encoded by the LCR16a core duplicon in the form of main and alternative ORFs, or noncoding features. A deeper understanding of any functional relationship between the NPIP main ORFs and the downstream PIPPI sORFs is pending a functional characterization of the NPIP protein family, albeit preliminary analysis suggests that NPIP proteins are upregulated by signals of viral infection, and hence may also participate in stress responses [50]. The entire NPIP gene family is undergoing rapid positive selection in primates [46, 48], suggesting that these evolutionarily young proteins could perform specialized functions in primate-specific tissues and organs.

For further experiments in our study, we worked with the exogenously expressed microprotein to define its localization and interactome to set the basis for further characterization of PIPPI. Interestingly, PIPPI appeared to be localized in the ER lumen, despite being devoid of canonical signal peptides or ER localization motifs. We hypothesized that this occurs thanks to the action of the trafficking protein TMED10, which can orchestrate the secretion of a spectrum of cargos upon facilitating their uptake in the ERGIC [67]. From here, PIPPI can be shuttled to the ER via its interaction with the PDI-homologue chaperone ERp44 [64]. Once in the ER, we observed that PIPPI has the ability to positively affect cellular response to ER stress, allowing cells to proliferate better, and partially escape apoptosis, under these conditions. Challenges remain, such as understanding how and when specific sORFs are used by an organism, especially in the case of recently evolved sORFs, whose expression and function could be extremely narrow by cell type or condition. As an attenuator of the ISR, PIPPI may serve a general pro-survival function that could be invoked by cells under specific stress conditions, which we have yet to fully understand. In summary, our study adds to our understanding of the functional potential of microproteins, and provides strategies to explore the vast and largely uncharacterized human sORF translatome.

Acknowledgements

We thank Prof. Iris Finkemeier and Paulina Heinkow for mass spectrometry service. We thank the BIC facility for enabling usage of their AiryScan microscopy, the Fernandez Capetillo group for allowing us to use their InCell microscope, the Bartek and Lemmens groups for access to the Nikon Eclipse Ti2. The virus used in the overexpression screen was produced and purified by the Virus Tech Core Facility of the Karolinska Institutet. Illustrations from bioicons.com were used in schemes and graphical abstract.

Author contributions: Conceptualization: L.L., S.J.E., G.L.R., and A.M.A. Methodology: L.L., C.N., D.S., M.H., G.L.R., A.M.A., K.R.S., I.P., J.L., R.M.M.B., and S.J.E. Investigation: L.L., G.L.R., A.M.A., A.S., M.H., L.B., D.S., K.R.S., C.N., M.H., R.S., I.P., and S.J.E. Formal analysis: L.L., G.L.R., A.M.A., K.R.S., D.H., I.P., R.M.M.B., and S.J.E. Writing—original draft: L.L. and S.J.E. Writing—review and editing: G.L.R., L.L., A.M.A., D.S., C.N., A.S., M.H., I.P., and S.J.E. Funding acquisition: S.J.E. Resources: S.J.E. and J.L. Supervision: S.J.E.

Supplementary data

Supplementary data is available at NAR online.

Funding

S.J.E. acknowledges funding by the IngaBritt och Arne Lundbergs Forskningsstiftelse, the Karolinska Institutet SFO for Molecular Biosciences, Vetenskapsrådet (2015-04815, 2020-04313), H2020 ERC Starting Grant (715024 RAPID), Åke Wibergs Stiftelse (M15-0275), Cancerfonden (22 2354 Pj), the Ming Wai Lau Center for Reparative Medicine, the Ragnar Söderbergs Stiftelse, Stiftelsen för Strategisk Forskning (FFL7); and the Knut och Alice Wallenbergs Stiftelse (2010-0215). A.M.A. was supported by a SciLifeLab RED postdoctoral fellowship. D.S. was supported by a Boehringer Ingelheim Fonds PhD fellowship. Sequencing analyses were processed on the Uppsala Multidisciplinary Center for Advanced Computational Science (UPPMAX), provided by the National Academic Infrastructure for Supercomputing in Sweden (NAISS) under projects NAISS 2023/6-19, NAISS 2023/22-84, SNIC 2022/6-14. NAISS is funded by the Swedish Research Council through grant agreement no. 2022-06725. J.L. and R.M.M.B. acknowledge funding by Cancerfonden (23 2819 Pj 01 H), Erling Perssons Stiftelse (22/9-2020), and Swedish Research Council (2019-04830). K.R.S. is funded by The Cancer Research Foundations of Radiumhemmet (241272). Funding to pay the Open Access publication charges for this article was provided by Karolinska Institutet.

Data availability

Mass spectrometry data were uploaded to ProteomeXchange via PRIDE. PRIDE submission PXD057806 contains DIA experiments from A375 total proteome samples (Fig. 2C and Supplementary Fig. S2F and G) acquired on a Thermo Orbitrap Exploris 480. PDX066409 contains PIPPI antibody IP-MS/MS experiments (Fig. 2D and E and Supplementary Fig. S2H) acquired on a Thermo Astral and Bruker timsTOF. PXD058567 contains A375 total proteome samples (Fig. 3F and G) and GFP-Trap IP-MS/MS experiments from A375 and U2OS cells (Fig. 4A–D). Miscellaneous source data is available on Mendeley Data (doi:10.17632/shyhmt4f2x.1).

References

1. Plaza S, Menschaert G, Payre F. In search of lost small peptides. *Annu Rev Cell Dev Biol* 2017;33:391–416. https://doi.org/10.1146/annurev-cellbio-100616-060516
2. Makarewich CA, Olson EN. Mining for micropeptides. *Trends Cell Biol* 2017;27:685–96. https://doi.org/10.1016/j.tcb.2017.04.006
3. Mudge JM, Ruiz-Orera J, Prensner JR *et al.* Standardized annotation of translated open reading frames. *Nat Biotechnol* 2022;40:994–9. https://doi.org/10.1038/s41587-022-01369-0
4. Denli AM, Narvaiza I, Kerman BE *et al.* Primate-specific ORF0 contributes to retrotransposon-mediated diversity. *Cell* 2015;163:583–93. https://doi.org/10.1016/j.cell.2015.09.025
5. Arnoult N, Correia A, Ma J *et al.* Regulation of DNA repair pathway choice in S and G2 phases by the NHEJ inhibitor CYREN. *Nature* 2017;549:548–52. https://doi.org/10.1038/nature24023
6. Slavoff SA, Heo J, Budnik BA *et al.* A human short open reading frame (sORF)-encoded polypeptide that stimulates DNA end joining. *J Biol Chem* 2014;289:10950–7. https://doi.org/10.1074/jbc.C113.533968
7. Guo B, Zhai D, Cabezas E *et al.* Humanin peptide suppresses apoptosis by interfering with Bax activation. *Nature* 2003;423:456–61. https://doi.org/10.1038/nature01627
8. Koh M, Ahmad I, Ko Y *et al.* A short ORF-encoded transcriptional regulator. *Proc Natl Acad Sci USA* 2021;118:e2021943118. https://doi.org/10.1073/pnas.2021943118
9. Lee C, Zeng J, Drew BG *et al.* The mitochondrial-derived peptide MOTS-c promotes metabolic homeostasis and reduces obesity and insulin resistance. *Cell Metab* 2015;21:443–54. https://doi.org/10.1016/j.cmet.2015.02.009
10. Makarewich CA, Baskin KK, Munir AZ *et al.* MOXI Is a mitochondrial micropeptide that enhances fatty acid β-oxidation. *Cell Rep* 2018;23:3701–9. https://doi.org/10.1016/j.celrep.2018.05.058

11. Pauli A, Norris ML, Valen E *et al.* Toddler: an embryonic signal that promotes cell movement via Apelin receptors. *Science* 2014:343:1248636. https://doi.org/10.1126/science.1248636
12. Chng SC, Ho L, Tian J *et al.* ELABELA: a hormone essential for heart development signals via the apelin receptor. *Dev Cell* 2013;27:672–80. https://doi.org/10.1016/j.devcel.2013.11.002
13. Tatemoto K, Hosoya M, Habata Y *et al.* Isolation and characterization of a novel endogenous peptide ligand for the human APJ receptor. *Biochem Biophys Res Commun* 1998;251:471–6. https://doi.org/10.1006/bbrc.1998.9489
14. Matsumoto A, Pasut A, Matsumoto M *et al.* mTORC1 and muscle regeneration are regulated by the LINC00961-encoded SPAR polypeptide. *Nature* 2017;541:228–32. https://doi.org/10.1038/nature21034
15. Schlesinger D, Elsässer SJ. Revisiting sORFs: overcoming challenges to identify and characterize functional microproteins. *FEBS J* 2022;289:53–74. https://doi.org/10.1111/febs.15769
16. Wright BW, Yi Z, Weissman JS *et al.* The dark proteome: translation from noncanonical open reading frames. *Trends Cell Biol* 2022;32:243–58. https://doi.org/10.1016/j.tcb.2021.10.010
17. Fulda S, Gorman AM, Hori O *et al.* Cellular stress responses: cell survival and cell death. *Int J Cell Biol* 2010;2010:214074.
18. Martinez TF, Chu Q, Donaldson C *et al.* Accurate annotation of human protein-coding small open reading frames. *Nat Chem Biol* 2020;16:458–68. https://doi.org/10.1038/s41589-019-0425-0
19. Andreev DE, O'Connor PBF, Fahey C *et al.* Translation of 5' leaders is pervasive in genes resistant to eIF2 repression. *eLife* 2015;4:e03971. https://doi.org/10.7554/eLife.03971
20. Fijalkowska D, Verbruggen S, Ndah E *et al.* eIF1 modulates the recognition of suboptimal translation initiation sites and steers gene expression via uORFs. *Nucleic Acids Res* 2017;45:7997–8013. https://doi.org/10.1093/nar/gkx469
21. Starck SR, Tsai JC, Chen K *et al.* Translation from the 5' untranslated region shapes the integrated stress response. *Science* 2016:351:aad3867. https://doi.org/10.1126/science.aad3867
22. Boumahdi S, de Sauvage FJ. The great escape: tumour cell plasticity in resistance to targeted therapy. *Nat Rev Drug Discov* 2020;19:39–56. https://doi.org/10.1038/s41573-019-0044-1
23. Khitun A, Ness TJ, Slavoff SA. Small open reading frames and cellular stress responses. *Mol Omics* 2019;15:108–16. https://doi.org/10.1039/C8MO00283E
24. Hofman DA, Prensner JR, van Heesch S. Microproteins in cancer: identification, biological functions, and clinical implications. *Trends Genet* 2025;41:146–61. https://doi.org/10.1016/j.tig.2024.09.002
25. Bray NL, Pimentel H, Melsted P *et al.* Near-optimal probabilistic RNA-seq quantification. *Nat Biotechnol* 2016;34:525–7. https://doi.org/10.1038/nbt.3519
26. Schindelin J, Arganda-Carreras I, Frise E *et al.* Fiji: an open-source platform for biological-image analysis. *Nat Methods* 2012;9:676–82. https://doi.org/10.1038/nmeth.2019
27. Stirling DR, Swain-Bowden MJ, Lucas AM *et al.* CellProfiler 4: improvements in speed, utility and usability. *BMC Bioinformatics* 2021;22:433. https://doi.org/10.1186/s12859-021-04344-9
28. Wickham H. *ggplot2 – Elegant Graphics for Data Analysis*. New York, NY: Springer-Verlag New York, 2016;
29. Team RS. *RStudio: Integrated Development Environment for R*. Boston, MA: PBC, 2022.
30. Cox J, Mann M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat Biotechnol* 2008;26:1367–72. https://doi.org/10.1038/nbt.1511
31. UniProt Consortium. Uniprot: the universal protein knowledgebase in 2023. *Nucleic Acids Res* 2023;51:D523–31. https://doi.org/10.1093/nar/gkac1052
32. Mellacheruvu D, Wright Z, Couzens AL *et al.* The CRAPome: a contaminant repository for affinity purification-mass spectrometry data. *Nat Methods* 2013;10:730–6. https://doi.org/10.1038/nmeth.2557
33. Zhang X, Smits AH, van Tilburg GB *et al.* Proteome-wide identification of ubiquitin interactions using UbIA-MS. *Nat Protoc* 2018;13:530–50. https://doi.org/10.1038/nprot.2017.147
34. Goedhart J, Luijsterburg MS. VolcaNoseR is a web app for creating, exploring, labeling and sharing volcano plots. *Sci Rep* 2020;10:20560. https://doi.org/10.1038/s41598-020-76603-3
35. Schlesinger D, Dirks C, Luzon CN *et al.* A large-scale sORF screen identifies putative microproteins involved in cancer cell fitness *iScience* 2025;28:111884. https://doi.org/10.1016/j.isci.2025.111884.
36. Karran P, Attard N. Thiopurines in current medical practice: molecular mechanisms and contributions to therapy-related cancer. *Nat Rev Cancer* 2008;8:24–36. https://doi.org/10.1038/nrc2292
37. Samandi S, Roy AV, Delcourt V *et al.* Deep transcriptome annotation enables the discovery and functional characterization of cryptic small proteins. *eLife* 2017;6:e27860. https://doi.org/10.7554/eLife.27860
38. Fornace AJ, Alamo I, Hollander MC. DNA damage-inducible transcripts in mammalian cells. *Proc Natl Acad Sci USA* 1988;85:8800–4. https://doi.org/10.1073/pnas.85.23.8800
39. Oyadomari S, Mori M. Roles of CHOP/GADD153 in endoplasmic reticulum stress. *Cell Death Differ* 2004;11:381–9. https://doi.org/10.1038/sj.cdd.4401373
40. Chen J, Brunner A-D, Cogan JZ *et al.* Pervasive functional translation of noncanonical human open reading frames. *Science* 2020;367:1140–6. https://doi.org/10.1126/science.aay0262
41. Jousse C, Bruhat A, Carraro V *et al.* 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. https://doi.org/10.1093/nar/29.21.4341
42. 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. https://doi.org/10.1074/jbc.M110.216093
43. Young SK, Palam LR, Wu C *et al.* Ribosome Elongation stall directs gene-specific translation in the integrated stress response. *J Biol Chem* 2016;291:6546–58. https://doi.org/10.1074/jbc.M115.705640
44. Lee H-C, Fu C-Y, Lin C-Y *et al.* Poly(U)-specific endoribonuclease ENDOU promotes translation of human CHOP mRNA by releasing uORF element-mediated inhibition. *EMBO J* 2021;40:e104123. https://doi.org/10.15252/embj.2019104123
45. Ma J, Ward CC, Jungreis I *et al.* Discovery of human sORF-encoded polypeptides (SEPs) in cell lines and tissue. *J Proteome Res* 2014;13:1757–65. https://doi.org/10.1021/pr401280w
46. Johnson ME, Viggiano L, Bailey JA *et al.* Positive selection of a gene family during the emergence of humans and African apes. *Nature* 2001;413:514–9. https://doi.org/10.1038/35097067
47. Cantsilieris S, Sunkin SM, Johnson ME *et al.* An evolutionary driver of interspersed segmental duplications in primates. *Genome Biol* 2020;21:202. https://doi.org/10.1186/s13059-020-02074-4
48. Dishuck PC, Munson KM, Lewis AP *et al.* Structural variation, selection, and diversification of the NPIP gene family from the human pangenome. *Cell Genomics* 2025.10:100977. https://doi.org/10.1016/j.xgen.2025.100977.
49. Martin FJ, Amode MR, Aneja A *et al.* Ensembl 2023. *Nucleic Acids Res* 2023;51:D933–41. https://doi.org/10.1093/nar/gkac958
50. Bekpen C, Baker C, Hebert MD *et al.* Functional characterization of the morpheus gene family. bioRxiv, https://doi.org/10.1101/116087, 12 March 2017, preprint: not peer reviewed.
51. Lafranchi L, Schlesinger D, Kimler KJ *et al.* Universal single-residue terminal labels for fluorescent live cell imaging of microproteins. *J Am Chem Soc* 2020;142:20080–7. https://doi.org/10.1021/jacs.0c09574

52. Breitling J, Aebi M. N-linked protein glycosylation in the endoplasmic reticulum. *Cold Spring Harb Perspect Biol* 2013;5:a013359. https://doi.org/10.1101/cshperspect.a013359
53. Bi K, Nishihara K, Machleidt T *et al.* Identification of known drugs targeting the endoplasmic reticulum stress response. *Anal Bioanal Chem* 2015;407:5343–51. https://doi.org/10.1007/s00216-015-8694-2
54. Slaine PD, Kleer M, Duguay BA *et al.* Thiopurines activate an antiviral unfolded protein response that blocks influenza A virus glycoprotein accumulation. *J Virol* 2021;95:e00453–21. https://doi.org/10.1128/JVI.00453-21
55. Hetz C, Zhang K, Kaufman RJ. Mechanisms, regulation and functions of the unfolded protein response. *Nat Rev Mol Cell Biol* 2020;21:421–38. https://doi.org/10.1038/s41580-020-0250-z
56. Daehn I, Brem R, Barkauskaite E *et al.* 6-Thioguanine damages mitochondrial DNA and causes mitochondrial dysfunction in human cells. *FEBS Lett* 2011;585:3941–6. https://doi.org/10.1016/j.febslet.2011.10.040
57. Yoshida T, Sakai T. Tunicamycin. In: Schwab M (ed.), *Encyclopedia of Cancer*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2011; pp. 3816–9. https://doi.org/10.1007/978-3-642-16483-5
58. Rutkowski DT, Kaufman RJ. A trip to the ER: coping with stress. *Trends Cell Biol* 2004;14:20–8. https://doi.org/10.1016/j.tcb.2003.11.001
59. Pakos-Zebrucka K, Koryga I, Mnich K *et al.* The integrated stress response. *EMBO Rep* 2016;17:1374–95. https://doi.org/10.15252/embr.201642195
60. Harding HP, Novoa I, Zhang Y *et al.* Regulated translation initiation controls stress-induced gene expression in mammalian cells. *Mol Cell* 2000;6:1099–108. https://doi.org/10.1016/S1097-2765(00)00108-8
61. Nougarède A, Tesnière C, Ylanko J *et al.* Improved IRE1 and PERK pathway sensors for multiplex endoplasmic reticulum stress assay reveal stress response to nuclear dyes used for image segmentation. *Assay Drug Dev Technol* 2018;16:350–60. https://doi.org/10.1089/adt.2018.862
62. Liu C, Li X, Li C *et al.* SLC3A2 is a novel endoplasmic reticulum stress-related signaling protein that regulates the unfolded protein response and apoptosis. *PLoS One* 2018;13:e0208993. https://doi.org/10.1371/journal.pone.0208993
63. Yang X, Xia R, Yue C *et al.* ATF4 regulates CD4+ T cell immune responses through metabolic reprogramming. *Cell Rep* 2018;23:1754–66. https://doi.org/10.1016/j.celrep.2018.04.032
64. Tempio T, Anelli T. The pivotal role of ERp44 in patrolling protein secretion. *J Cell Sci* 2020;133:jcs240366. https://doi.org/10.1242/jcs.240366
65. Higo T, Hattori M, Nakamura T *et al.* Subtype-specific and ER lumenal environment-dependent regulation of inositol 1,4,5-trisphosphate receptor type 1 by ERp44. *Cell* 2005;120:85–98. https://doi.org/10.1016/j.cell.2004.11.048
66. Bremser M, Nickel W, Schweikert M *et al.* Coupling of coat assembly and vesicle budding to packaging of putative cargo receptors. *Cell* 1999;96:495–506. https://doi.org/10.1016/S0092-8674(00)80654-6
67. Zhang M, Liu L, Lin X *et al.* A translocation pathway for vesicle-mediated unconventional protein secretion. *Cell* 2020;181:637–652.e15. https://doi.org/10.1016/j.cell.2020.03.031
68. Oka OBV, Pringle MA, Schopp IM *et al.* ERdj5 is the ER reductase that catalyzes the removal of non-native disulfides and correct folding of the LDL receptor. *Mol Cell* 2013;50:793–804. https://doi.org/10.1016/j.molcel.2013.05.014
69. Ushioda R, Hoseki J, Araki K *et al.* ERdj5 is required as a disulfide reductase for degradation of misfolded proteins in the ER. *Science* 2008;321:569–72. https://doi.org/10.1126/science.1159293
70. Anelli T, Alessio M, Bachi A *et al.* Thiol-mediated protein retention in the endoplasmic reticulum: the role of ERp44. *EMBO J* 2003;22:5015–22. https://doi.org/10.1093/emboj/cdg491
71. Bekpen C, Tautz D. Human core duplicon gene families: game changers or game players? *Brief Funct Genomics* 2019;18:402–11. https://doi.org/10.1093/bfgp/elz016

Received: December 8, 2024. **Revised:** September 21, 2025. **Accepted:** September 25, 2025