Supplementary Materials for

Increasing cell size remodels the proteome and promotes senescence

Michael C. Lanz†, Evgeny Zatulovskiy†*, Matthew P. Swaffer, Lichao Zhang, Ilayda Ilerten, Shuyuan Zhang, Dong Shin You, Georgi Marinov, Patrick McAlpine, Josh E. Elias, Jan M. Skotheim*

† These authors contributed equally to this work
* Corresponding authors: Email: skotheim@stanford.edu ; evgeny@stanford.edu

This PDF file includes:

Figure S1 to S15

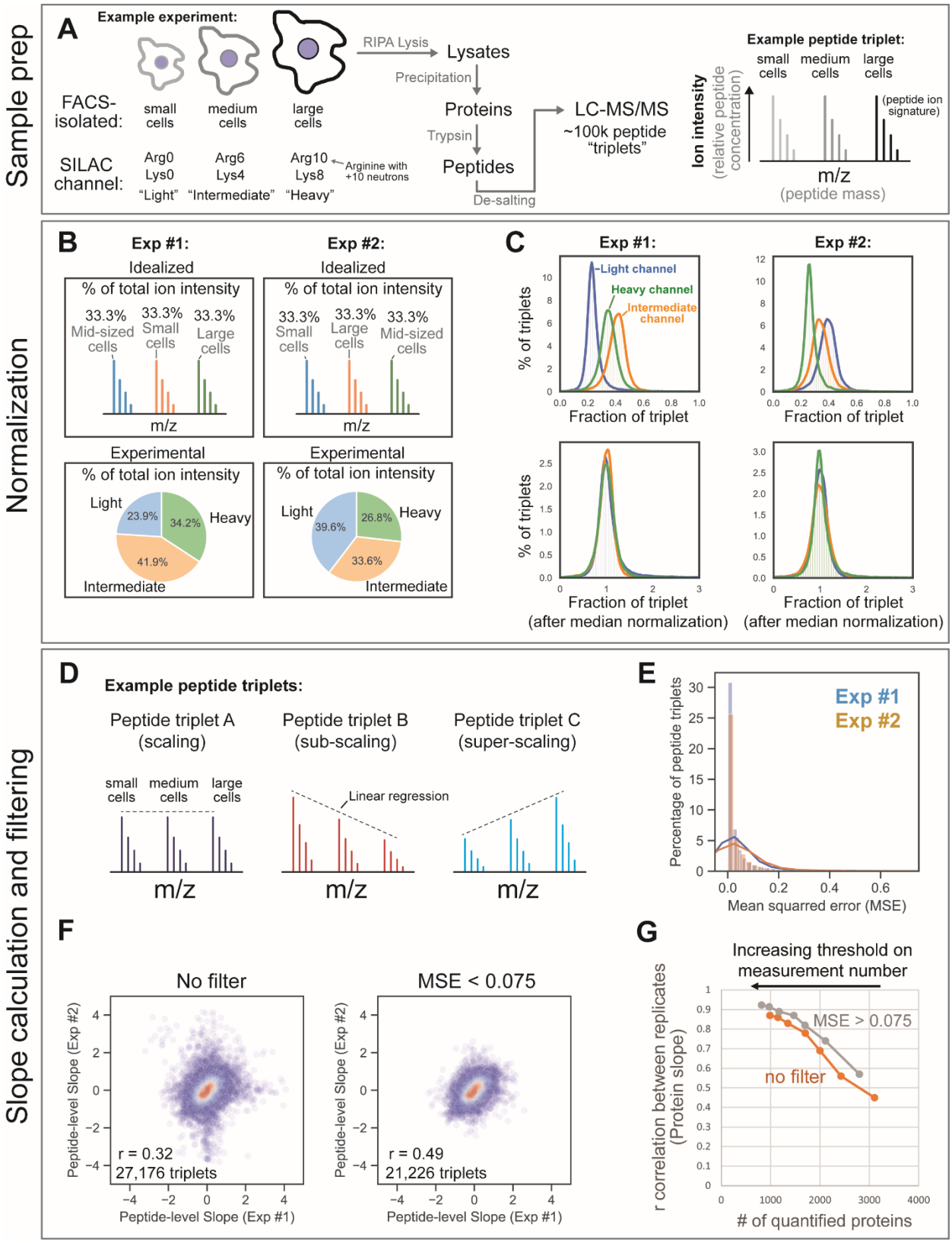
Sample prep
A
Example experiment:
FACS-isolated:
small cells
medium cells
large cells
SILAC channel:
Arg0 Lys0
Arg6 Lys4
Arg10 Lys8
Arginine with +10 neutrons
"Light"
"Intermediate"
"Heavy"
RIPA Lysis
Lysates
Precipitation
Proteins
Trypsin
Peptides
De-salting
LC-MS/MS
~100k peptide "triplets"
Example peptide triplet:
small cells
medium cells
large cells
Ion intensity
(relative peptide concentration)
(peptide ion signature)
m/z
(peptide mass)
Normalization
B
Exp #1:
Exp #2:
Idealized
% of total ion intensity
33.3% Mid-sized cells
33.3% Small cells
33.3% Large cells
33.3% Small cells
33.3% Large cells
33.3% Mid-sized cells
m/z
Experimental
% of total ion intensity
Light
Heavy
Intermediate
23.9%
34.2%
41.9%
39.6%
26.8%
33.6%
C
Exp #1:
Exp #2:
Light channel
Heavy channel
Intermediate channel
% of triplets
Fraction of triplet
Fraction of triplet (after median normalization)
Slope calculation and filtering
D
Example peptide triplets:
Peptide triplet A (scaling)
Peptide triplet B (sub-scaling)
Peptide triplet C (super-scaling)
small cells
medium cells
large cells
Linear regression
m/z
E
Exp #1
Exp #2
Percentage of peptide triplets
Mean squarred error (MSE)
F
No filter
MSE < 0.075
Peptide-level Slope (Exp #2)
Peptide-level Slope (Exp #1)
r = 0.32
27,176 triplets
r = 0.49
21,226 triplets
G
Increasing threshold on measurement number
MSE > 0.075
no filter
r correlation between replicates (Protein slope)
# of quantified proteins

Figure S1. A SILAC proteomics method to measure how the proteome changes as a function of G1 cell size, related to Figure 1: (A) Human cells were metabolically labeled in cell culture, sorted by G1 cell size using FACS, and subjected to proteomic analysis. **(B)** Differences in the amount of total protein contributed from the small-, medium-, or large-cell size populations were normalized using the signal proportionality from the light, intermediate, and heavy channel of each peptide triplet. Small, medium, and large cells were mixed prior to lysis, so the amount of protein in each SILAC channel is uneven. Rather than normalize L/H and L/M SILAC ratios separately, we normalize all three channels together so that the values in our dataset represent relative changes to each peptide. **(C)** For each individual peptide triplet, we determined the fraction of the triplet's total ion intensity present in each SILAC channel. The distributions of these fractions were then adjusted by the median (see methods for a complete description of the normalization process). **(D)** Peptide slope values are calculated from a linear regression of the relative ion intensity in each SILAC channel and mean cell size. Mean cell size was determined by Coulter counter prior to mixing and lysis. **(E)** Distribution of mean squared error values for peptide triplet regressions (~50,000 per experiment). The mean squared error was used to track the linear fit of each peptide regression. **(F)** Correlation of peptide slopes calculated from biological replicate experiments before and after applying a filter for mean squared error (MSE). 27,176 unique peptide measurements were identified in both replicate experiments. A unique peptide measurement is defined by the peptide sequence, modification state, charge state, and the fraction number (the fraction number is only considered if the experiment was pre-fractioned and multiple fractions were analyzed). **(G)** Filtering peptides by mean squared error from a linear fit improves data quality. MSE filtering improves the correlation of protein slope values derived from biological

replicates, and the improvement is consistent across different thresholds of measurement confidence (*i.e.*, peptide measurements per protein).

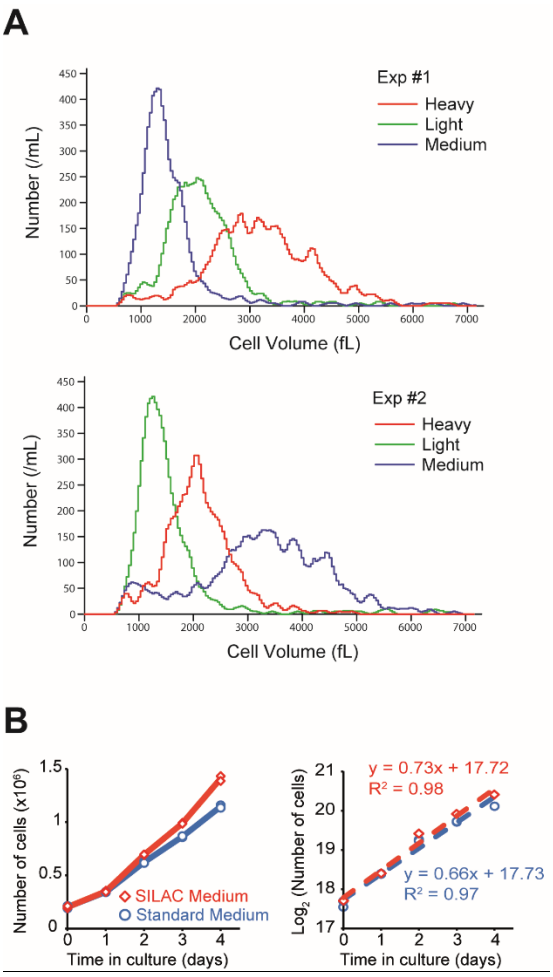

Figure S2. Supporting information for SILAC proteomics, related to Figure 1: (A) Coulter counter measurement of HLF cells isolated by FACS. Small, medium, and large cell populations are colored based on the SILAC labeling orientation for the two replicate experiments in Figure 1. See Data S3 for cell size measurements for all proteomic experiments. **(B)** HFL primary cell proliferation rates in SILAC vs standard medium.

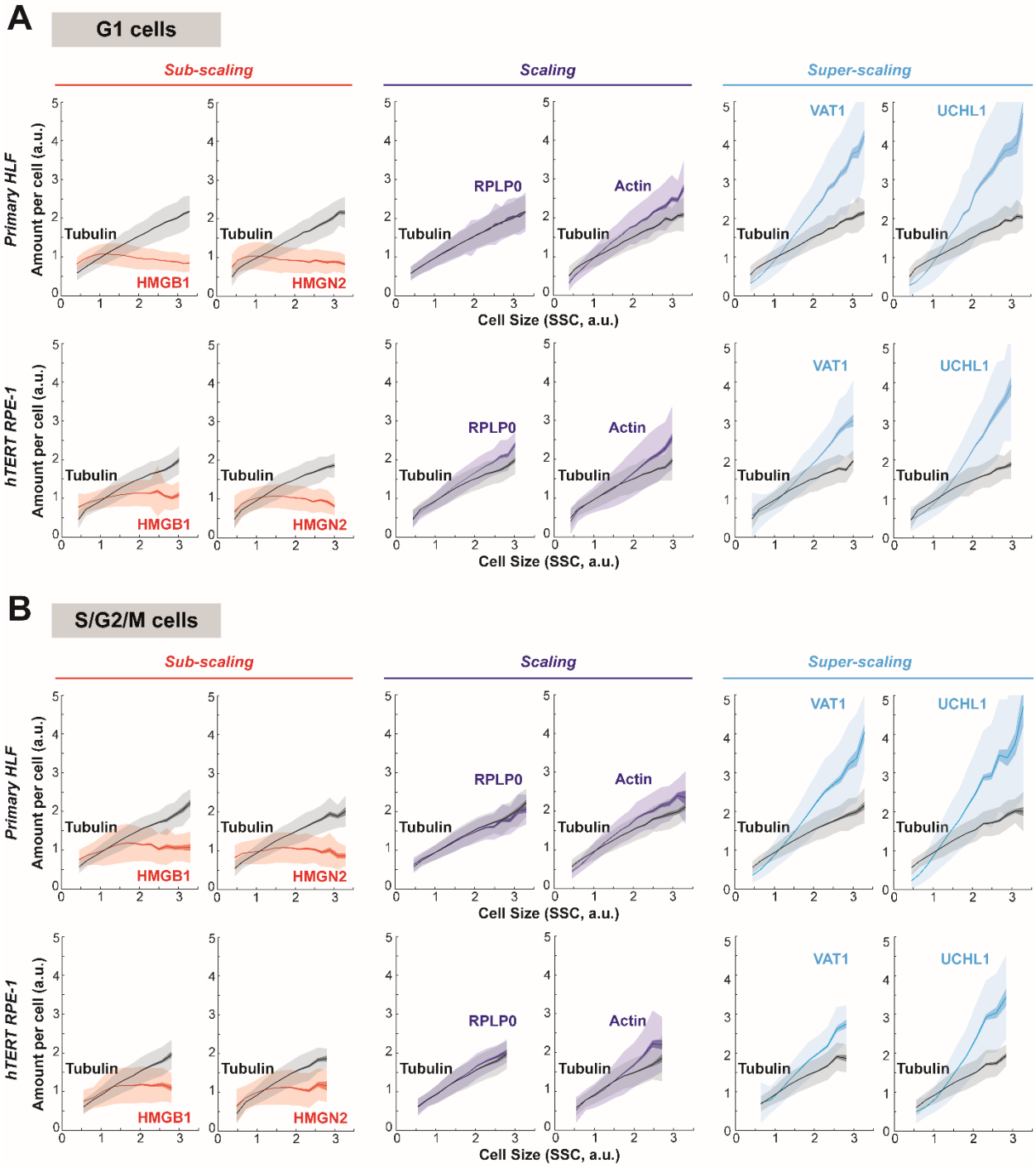

Figure S3. Validation of protein size-scaling behaviors using flow cytometry, related to Figure 1: Cycling HLF and RPE-1 cells were fixed and stained with antibodies against subscaling (HMGB1, HMGN2), scaling (RPLP0, beta-Actin), and superscaling proteins (VAT1, UCHL1).

Alpha-tubulin is used as an internal control for each sample. Using flow cytometry, G1 and non-G1 cells were gated by DNA content (DAPI dye) and analyzed separately in panels **(A)** and **(B)**, respectively. The data were binned by cell size (SSC, the side scatter parameter) and plotted as mean protein amounts per cell for each size bin (solid lines). Dark shaded area shows standard error of the mean for each bin, and light shaded area shows the standard deviation. A representative is shown of n=3 biological replicates for each experiment. 100,000 cells were analyzed for each sample.

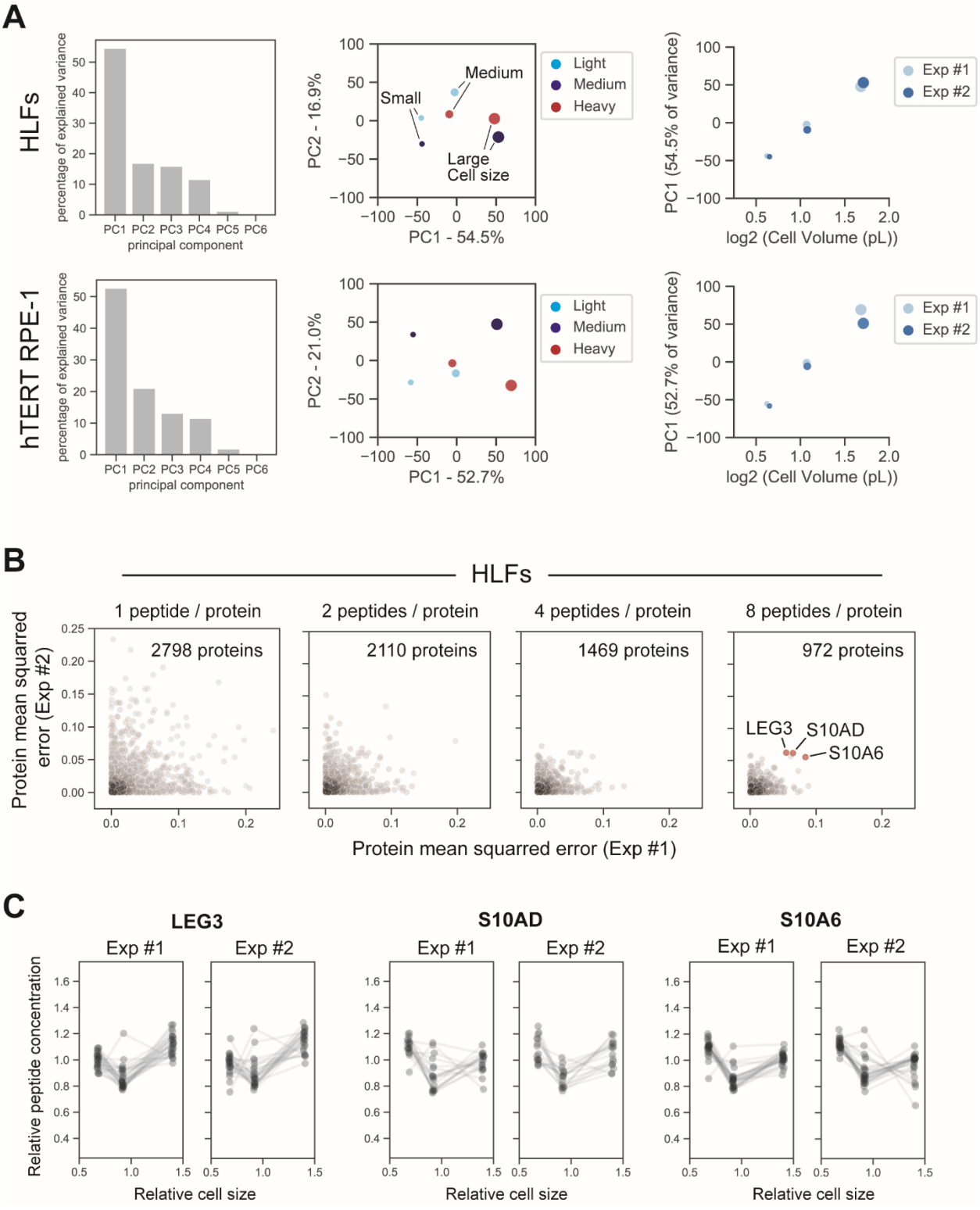

Figure S4. Cell size-dependent changes to concentrations in the proteome are mostly linear, related to Figure 1: (A) Principal component analysis (PCA) of the replicate experiments

performed using HLF and hTERT RPE-1 cell lines. Data frame input for the PCA contained the relative SILAC ion channel intensity ("light", "medium", and "heavy") for every measured protein in each experiment (after filtering for MSE). Dot size represents the mean cell size corresponding to each SILAC channel. PC1 represents the majority of variance in both experiments and correlates with the change in cell size. **(B)** Correlation for the mean squared error (MSE) of the Protein Slope regression from two biological replicates. A threshold for the minimum number of peptide measurements per protein is increased from left to right. Because very few proteins have reproducible large MSE values, we conclude that most proteins scale linearly with G1 cell size. **(C)** Peptide-level measurements for the few proteins with non-linear scaling are plotted below.

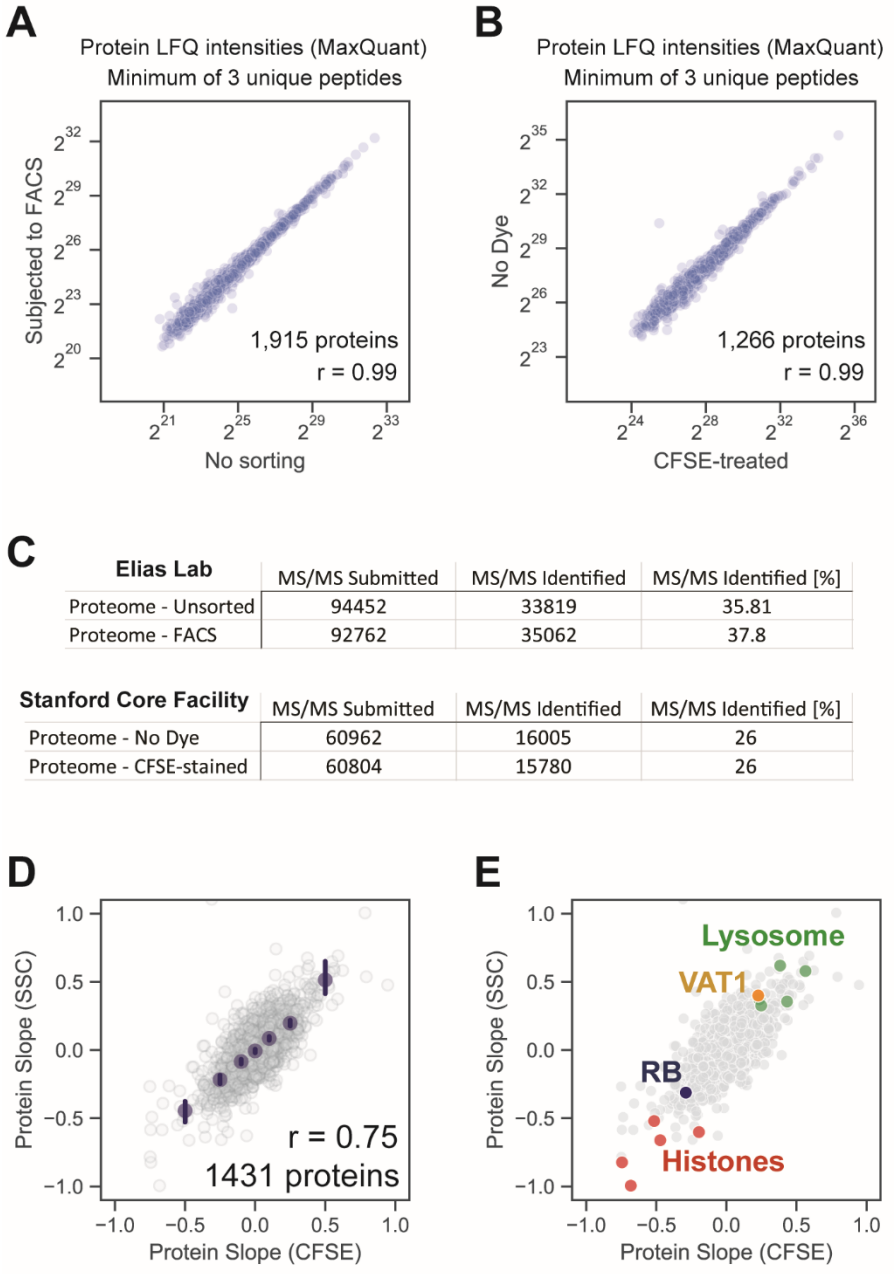

Figure S5. Controls indicating that cell sorting does not affect proteomics measurements, related to Figure 1: (A) A dish of HLF cells were split into two equal parts, and one half was run through the FACS machine while the other half sat on ice. MaxQuant LFQ was used to determine whether cells subjected to FACS exhibited altered proteomes. The strong correlation between the proteomes of sorted and unsorted cells suggests that FACS did not appreciably bias our

measurements. **(B)** MaxQuant LFQ was used to compare proteome samples from CFSE-stained and unlabeled cells. The strong correlation between the proteomes of CFSE-treated and untreated cells suggests that using a total protein dye does not appreciably bias our measurement. **(C)** Peptide discovery is not impacted by FACS or CFSE staining. **(D)** RPE-1 cells with G1 DNA content were sorted using total protein / cell (CFSE stain) or side scatter to achieve 3 bins of different sized cells. Protein Slope values derived from cells sorted by total protein and side scatter are compared. A select set of proteins from the comparison in (D) are highlighted **(E)**.

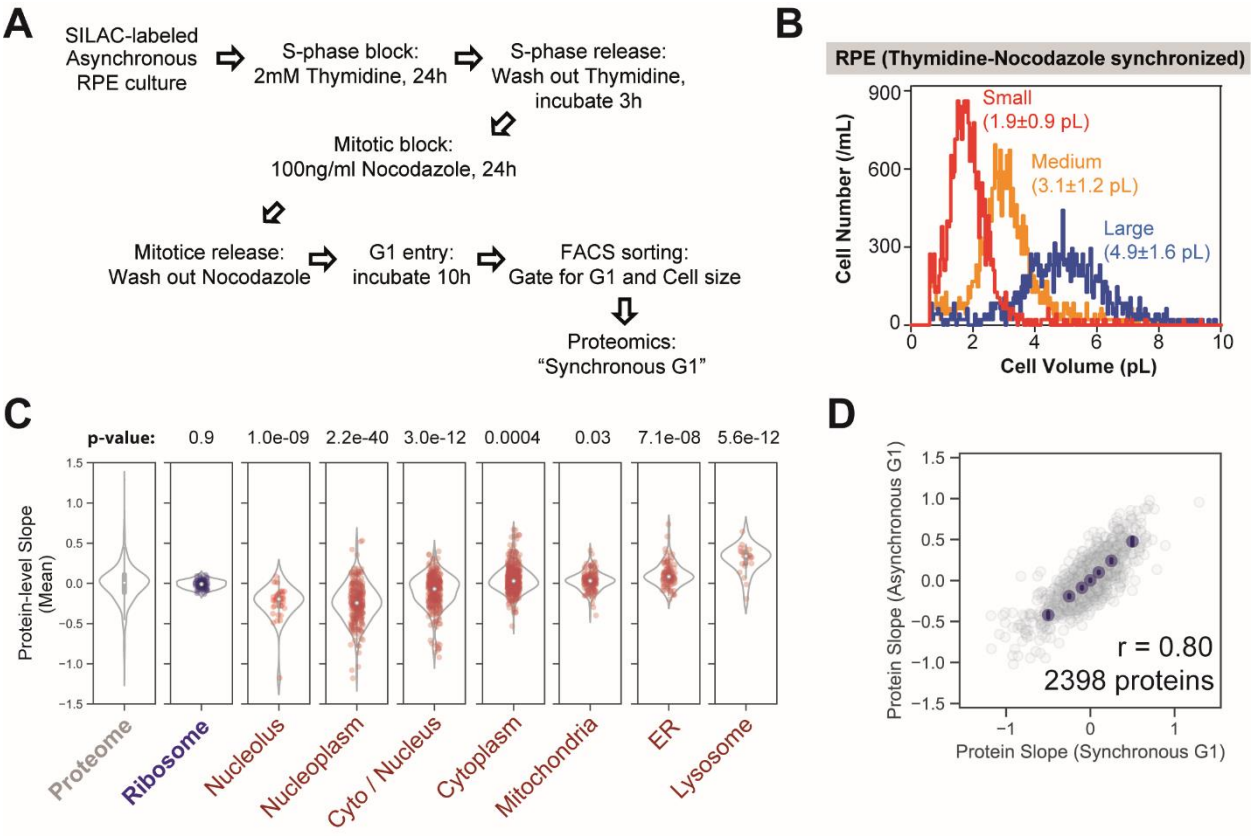

Figure S6. Changes to the proteome are primarily driven by cell size and not cell age, related to Figure 2: (A) Metabolically labeled RPE-1 cells were synchronously released into the cell cycle following a Thymidine-Nocodazole cell cycle arrest. **(B)** A similar distribution of G1 sizes were isolated from cells synchronously released into G1 and from asynchronous cultures. **(C)** Protein slopes were calculated as described in Figure 1. Only proteins with at least 4 peptide measurements in both replicate experiments are considered for the violin plots (Data S3). **(D)** Correlation of Protein Slope values calculated from RPE-1 cells synchronously released into G1 and from asynchronous cultures.

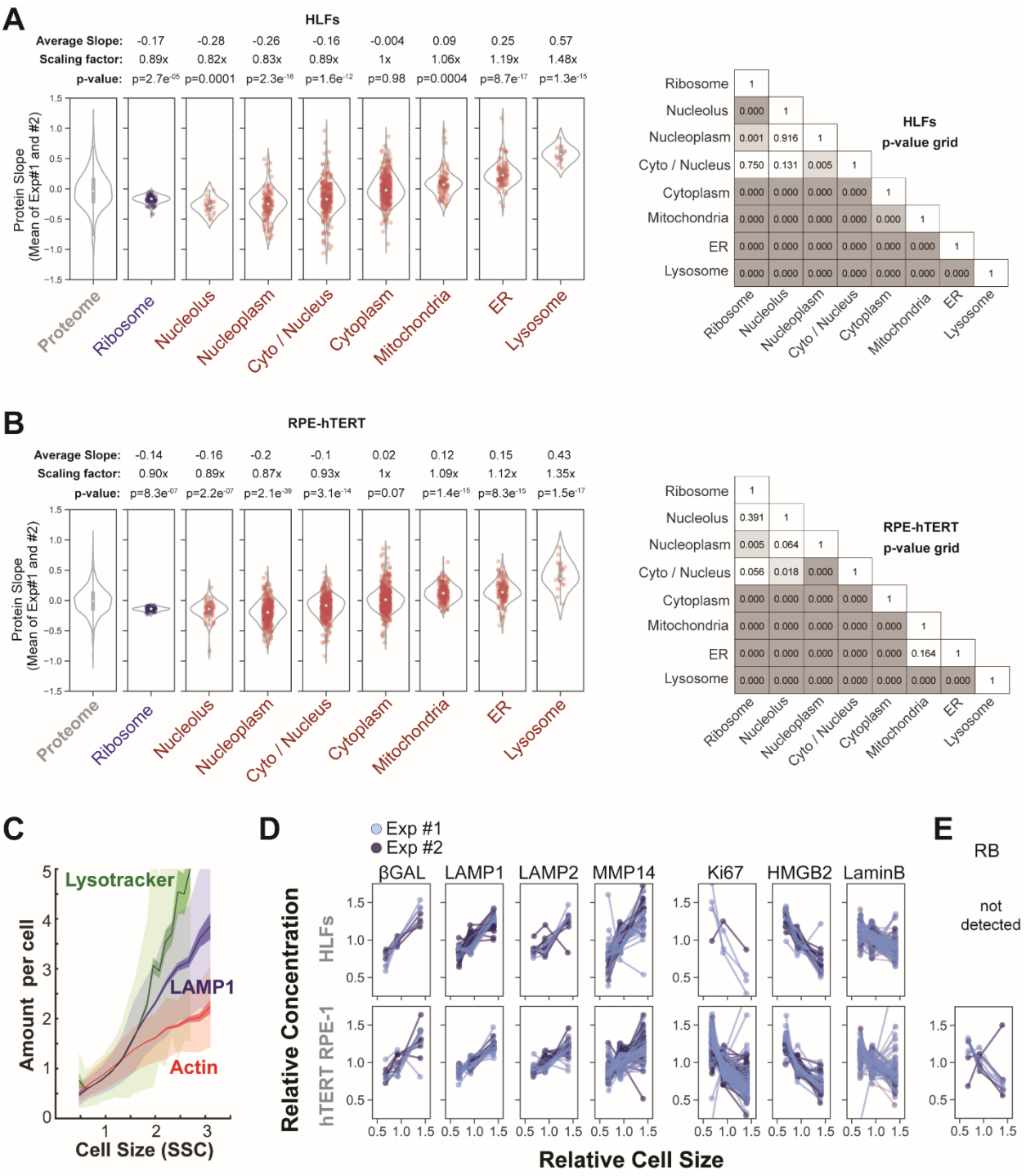

Figure S7. Size scaling of proteome content in an epithelial (RPE-1) and a fibroblast (HLFs) cell line, related to Figure 2: (A) Distribution of slopes derived from HLF cells for proteins associated with the indicated compartments. Violin plots depict the average slopes for the proteins

highlighted in Figure 2b. P-values above the violin plots are derived from a t-test between the indicated protein group and the rest of the dataset. t-tests comparing the slopes for each group of proteins are visualized in a grid format. **(B)** Replicate experiment using the immortalized RPE-1 cell line was performed as in (A). **(C)** Validation of lysosome super-scaling with cell size using flow cytometry. Both the lysosomal protein LAMP1 and the Lysotracker dye amount increase with cell size faster than Actin, which is a proxy for total protein. The data for G1 RPE-1 cells were binned by cell size (SSC, the side scatter parameter) and plotted as mean protein amounts per cell for each size bin (solid lines). Dark shaded area shows standard error of the mean for each bin, and light shaded area shows the standard deviation. A representative is shown of n=5 biological replicates for each experiment. About 100,000 cells were analyzed for each sample. **(D)** Examples from our proteomics data set of cell-size-dependent protein concentration changes in proliferating cells that are normally associated with senescence. **(E)** RB is diluted with increasing G1 cell size.

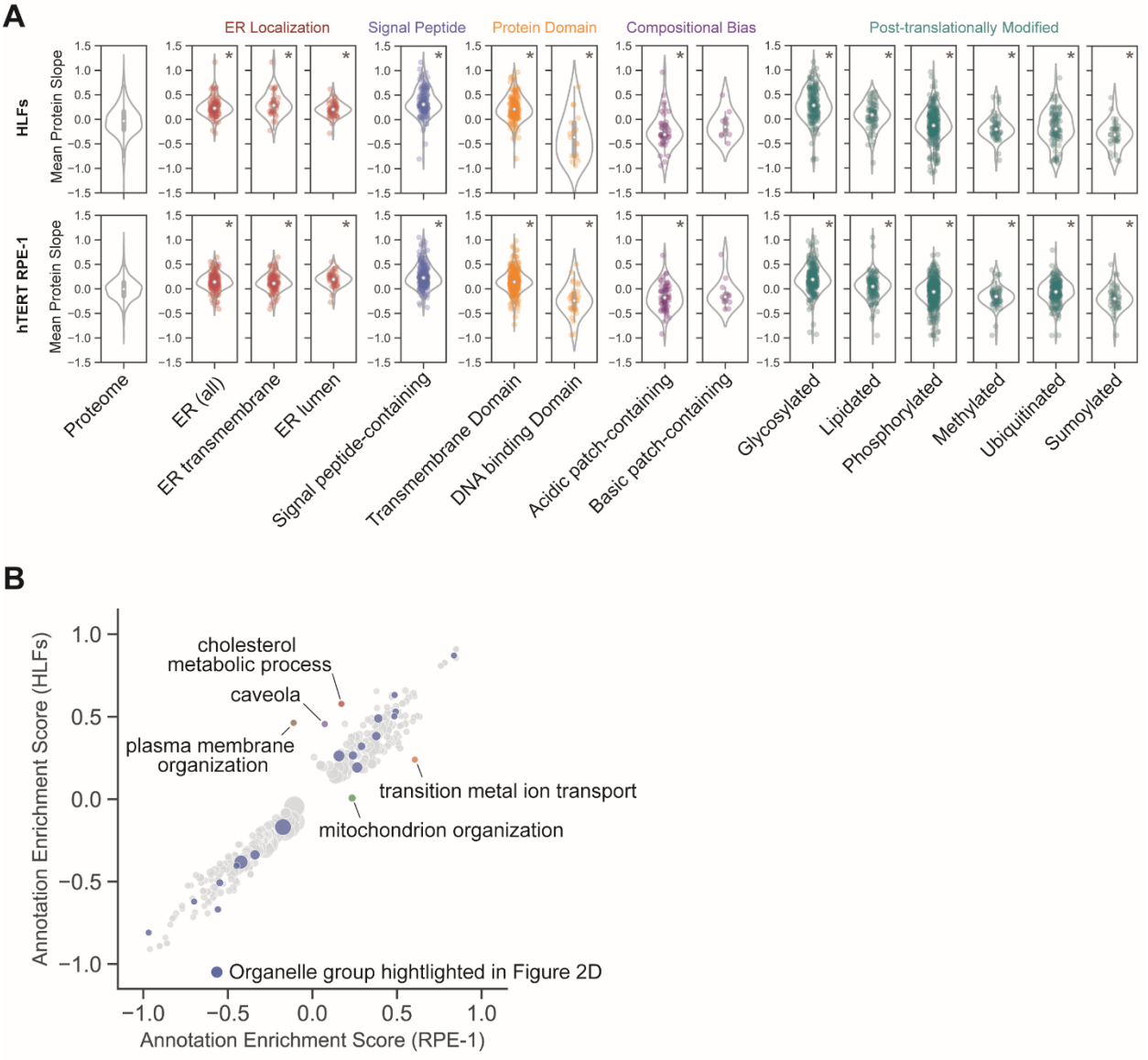

Figure S8. Bioinformatic analysis of scaling behavior, related to Figure 3: (A) Size scaling behaviors of various protein groups. Proteins were grouped by amino acid sequence features or the presence of post-translational modifications. Transmembrane and luminal ER proteins were differentiated by the presence or absence of a transmembrane domain, respectively. All groups were compared to the whole proteome via T-test (asterisk denotes a p-value < 0.05). **(B)** 2D annotation enrichment comparison of Protein Slope values calculated from hTERT RPE-1 and

HLF cell lines. Data S5 contains the Enrichment Scores for all annotation groups depicted in the plot.

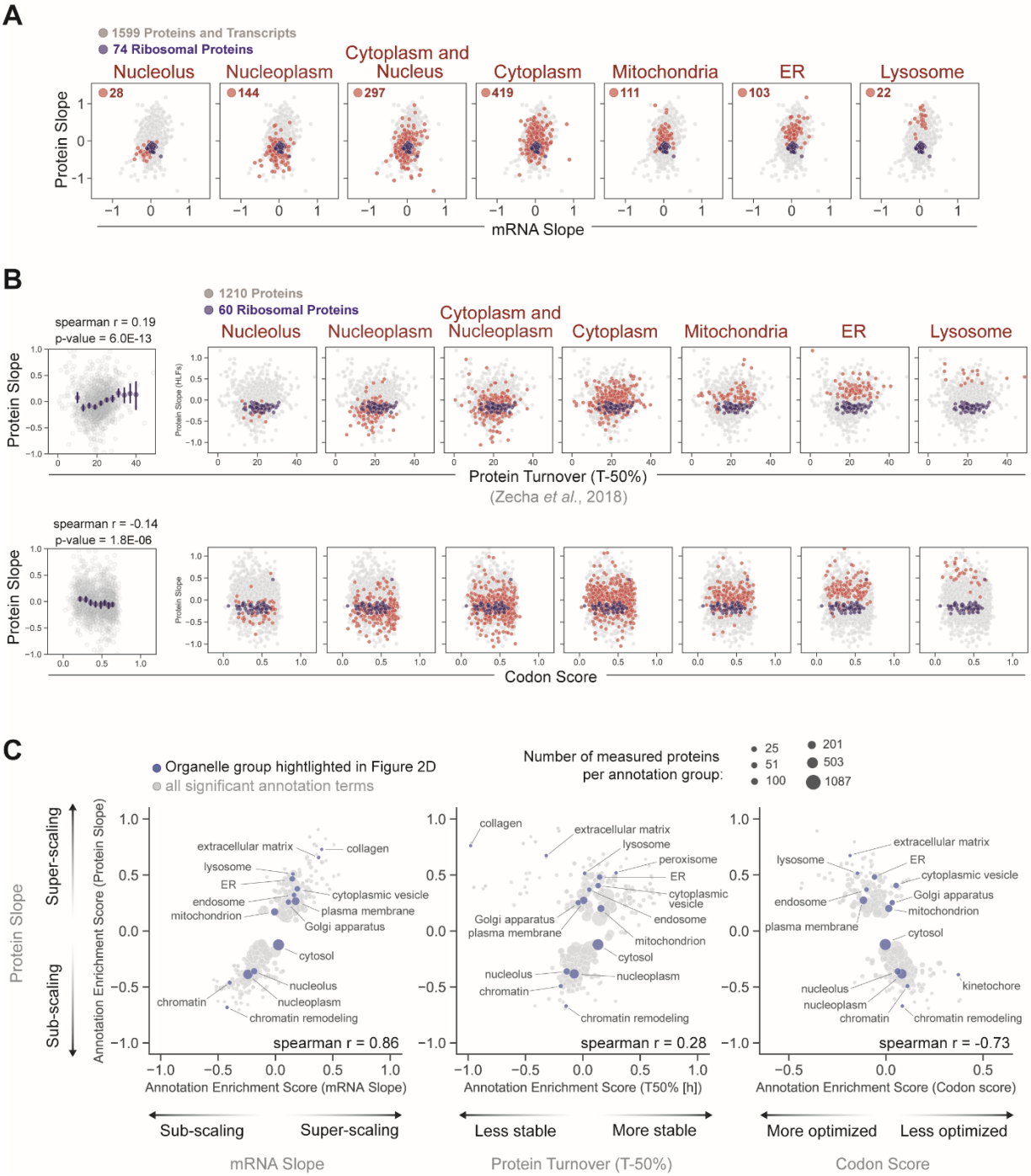

Figure S9. Transcriptional and post-transcriptional features correlate with the size scaling of individual proteins, related to Figure 3: (A) RNA Slope vs Protein Slope comparison (as in Figure 3). Proteins are highlighted based on their subcellular localization annotation described in Figure 2C. **(B)** Protein Slope comparison with two post-transcriptional features. Protein turnover was determined by pulse SILAC labeling (Zecha et al., 2018). Codon affinity is the percentage of codons in a given protein's mRNA transcript that contain an A or T base in the 3rd position. Only the 1,700 proteins included in the linear model were considered in these plots, as well as an addition requirement of >3 peptide per protein (1,210 proteins in total). **(C)** 2D annotation enrichment analysis for all components included in the linear model (Figure 3D). Data S8 contains the Enrichment Scores for all annotation groups depicted in the plot.

OLS Regression Results - Summarized

1,700 proteins common to Protein Slope, RNA Slope, and Protein Turnover (T-50%) datasets

Dependent variable - Protein Slope (HLF)

Independent variables	coef	std_err	t-stat	P>|t|	[0.025]	[0.975]			Prediction %
constant	-0.0082	0.005	-1.586	0.113	-0.018	0.002	R-squared	0.528	100%
Replicate #1 (vs #2)	0.689	0.016	43.552	0	0.658	0.72	Adj. R-squared	0.527	
							Prob (F-statistic)	7.53E-279	
constant	-0.025	0.007	-3.607	0	-0.038	0.011	R-squared	0.076	14%
RNA Slope	0.082	0.007	11.848	0	0.069	0.096	Adj. R-squared	0.076	
							Prob (F-statistic)	3.59E-31	
constant	-0.028	0.07	-4.101	0	-0.041	-0.041	R-squared	0.114	22%
Protein Turnover (T-50%) - Zecha *et al.*, 2018	0.042	0.07	5.928	0	0.028	0.028	Adj. R-squared	0.112	
Codon Score	-0.039	0.07	-5.765	0	-0.052	-0.026	Prob (F-statistic)	4.43E-44	
RNA Slope	0.083	0.07	12.206	0	0.070	0.097			
constant	0.032	0.009	3.423	0	0.014	0.051	R-squared	0.266	50%
Mitochondria_Localization	0.049	0.019	2.543	0.011	0.011	0.086	Adj. R-squared	0.264	
Lysosome_Localization	0.424	0.041	10.274	0	0.343	0.5	Prob (F-statistic)	4.07E-111	
ER_Localization	0.171	0.02	8.437	0	0.13	0.211			
Nucleus_Localization	-0.223	0.013	-16.887	0	-0.25	-0.197			
Ribosome_Localization	-0.190	0.033	-5.678	0	-0.255	-0.124			
constant	0.12	0.029	4.247	0	0.067	0.183	R-squared	0.314	60%
Protein Turnover (T-50%) - - Zecha *et al.*, 2018	0.087	0.020	4.360	0	0.048	0.13	Adj. R-squared	0.311	
Codon Score	-0.084	0.019	-4.532	0	-0.121	-0.046	Prob (F-statistic)	8.56E-133	
RNA Slope	0.183	0.020	9.274	0	0.145	0.222			
Mitochondria_Localization	0.193	0.059	3.252	0.001	0.077	0.310			
Lysosome_Localization	1.318	0.128	10.326	0	1.068	1.568			
ER_Localization	0.518	0.063	8.240	0	0.395	0.642			
Nucleus_Localization	-0.587	0.042	-13.821	0	-0.670	-0.504			
Ribosome_Localization	-0.609	0.103	-5.894	0	-0.812	-0.406			

Figure S10. Linear regression analysis predicts size scaling behavior of individual proteins, related to Figure 3: The prediction of size scaling behavior was based on the 1,700 proteins that are in the published protein turnover dataset (HeLa cells) (Zecha et al., 2018), as well as our RNA Slope, and Protein Slope datasets (at least 2 peptides / protein) that we report here. Independent variables for codon affinity, RNA Slope, and Protein turnover (time to replace 50% of a given protein species) were each independently standardized by subtracting all values by the dataset's mean and then dividing by the dataset's standard deviation. The subcellular localization variable was based on UniProt's "Subcellular location [CC]" annotations and entered as a binary value for each compartment (1 if a protein possessed an annotation and 0 if it did not). Only subcellular compartments that provided nonredundant predictive power were ultimately included in the model. A constant value was added to the regression equation using the add_constant function in statsmodels. We set the benchmark for predictive accuracy (Prediction %) as the correlation between biological replicate experiments, *i.e.*, Protein Slope from Exp #1 vs Exp #2.

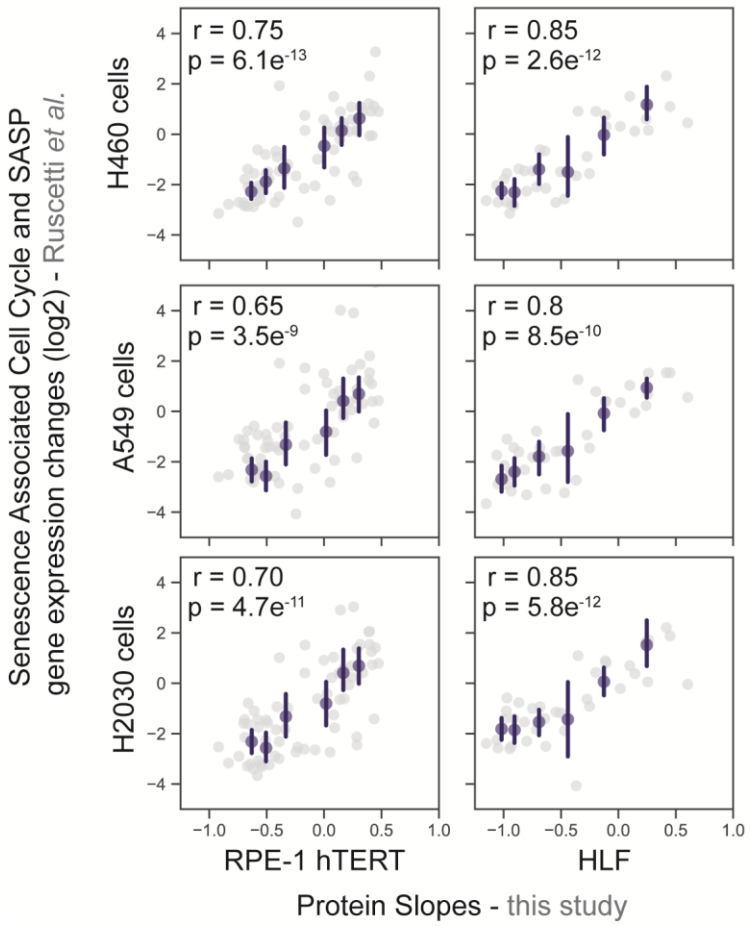

Figure S11. Senescence-associated proteome changes in proliferative HLF and hTERT RPE-1 cells, related to Figure 4: Size-dependent proteome changes from this study correlate with the senescence-associated SASP and cell cycle gene expression changes defined by Ruscetti, *et al.* (Ruscetti et al., 2018).

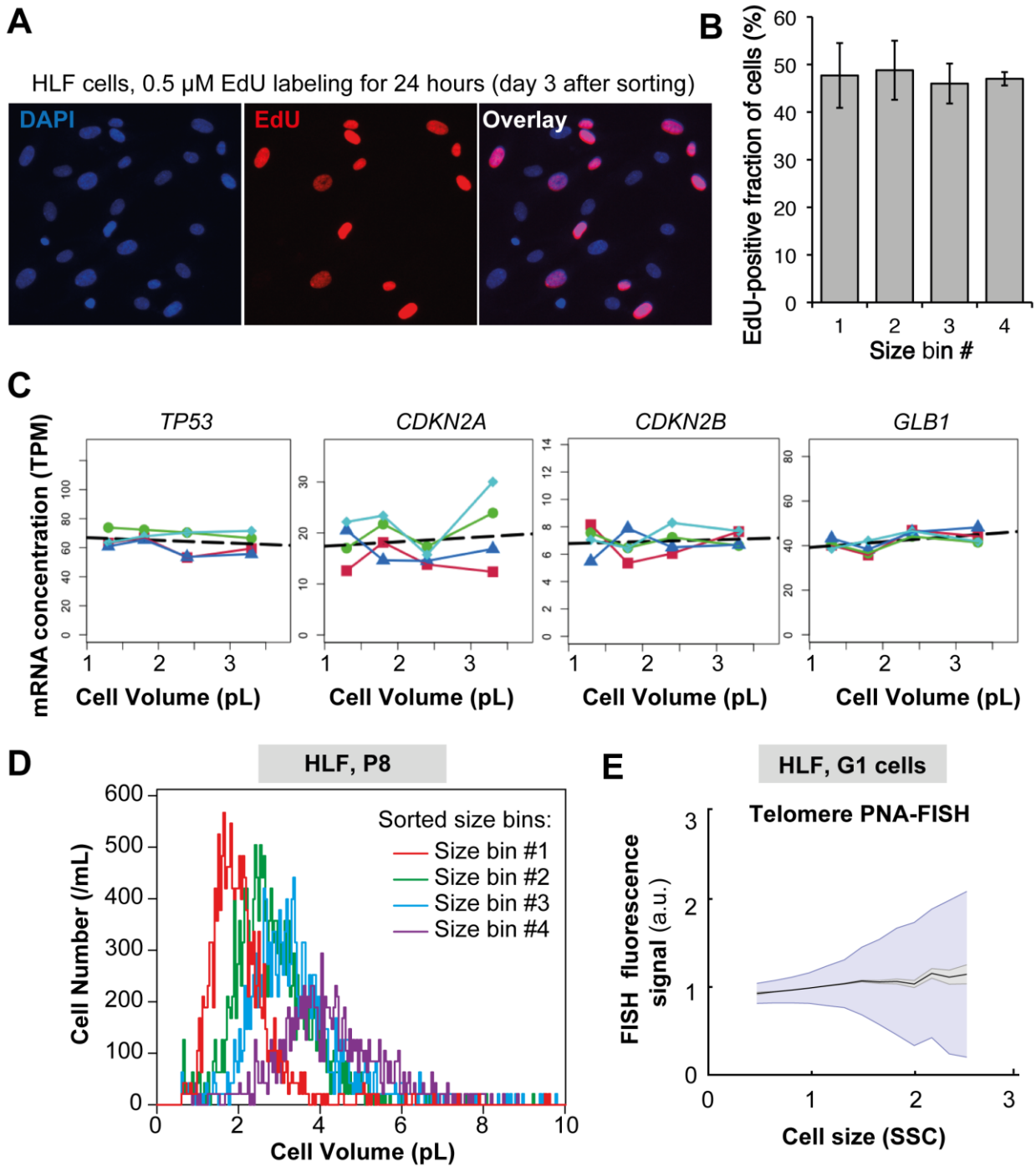

Figure S12. Analysis of proliferation and senescence in different-sized cells, related to Figure 4: (**A**) A representative image showing DAPI and EdU staining of HLF cells 3 days after size sorting. To evaluate the percentage of proliferating cells in each size bin after FACS sorting, the

sorted HLF cells were re-plated, allowed to settle for 3 days, and then incubated with EdU for 24 hours to label all the cells that enter S-phase within this period. EdU was detected using a Click-iT kit. (**B**) Percentage of EdU positive cells in each of the four cell size bins, 3 days after FACS sorting by size, as in (A). The error bars show s.e.m. n = 3 biological replicates. **(C)** Transcript levels of key senescence reporter genes in size-sorted HLF cells. G1 HLF cells were sorted into four size bins using FACS. The concentrations of *TP53, CDKN2A, CDKN2B,* and *GLB1* mRNAs were then determined by RNAseq and plotted against the mean cell size for each bin. Each colored line represents one of four replicates. **(D)** A representative example of cell size distributions for HLF cells sorted into 4 size bins by FACS and measured on a Coulter counter. (**E**) Flow-FISH Telomere peptide nucleic acid staining in passage #8 HLF cells plotted against cell size. The signal intensity, plotted on vertical axis, reflects the telomere length. The black line shows mean values for each size bin. The grey area is s.e.m. and the blue area shows standard deviation.

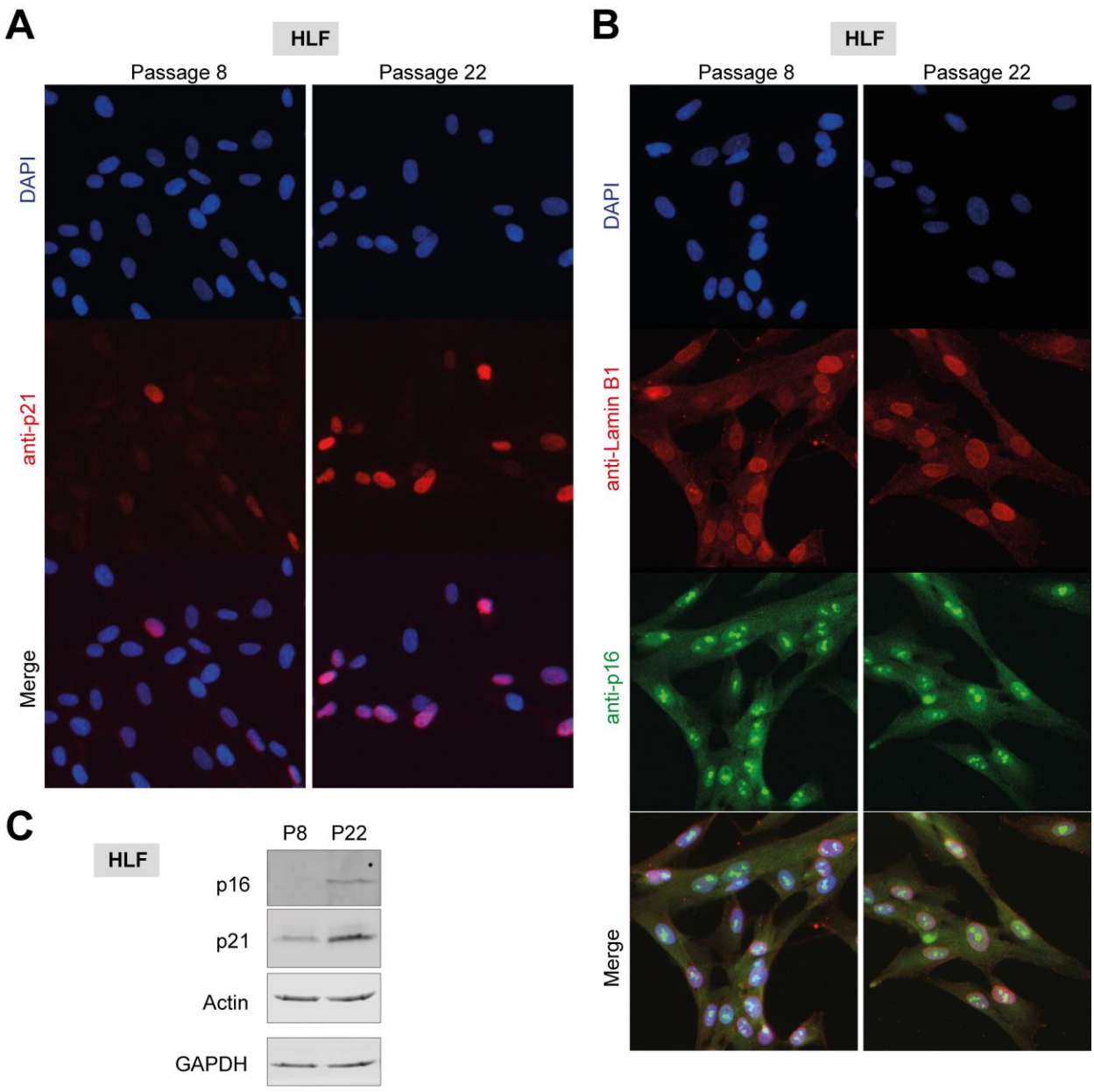

Figure S13. Validation of p21 and p16 antibodies for the detection of senescent cells, related to Figure 5: (**A-B**) Immunofluorescent staining of HLF cells at Passage 8 (non-senescent) and Passage 22 (undergoing senescence) with antibodies against p21 (A), Lamin B1, and p16 (B). DAPI staining labels cell nuclei. (**C**) Immunoblotting of HLF cell lysates at Passages 8 and 22 with antibodies against p16 and p21. Actin and GAPDH were used as loading controls.

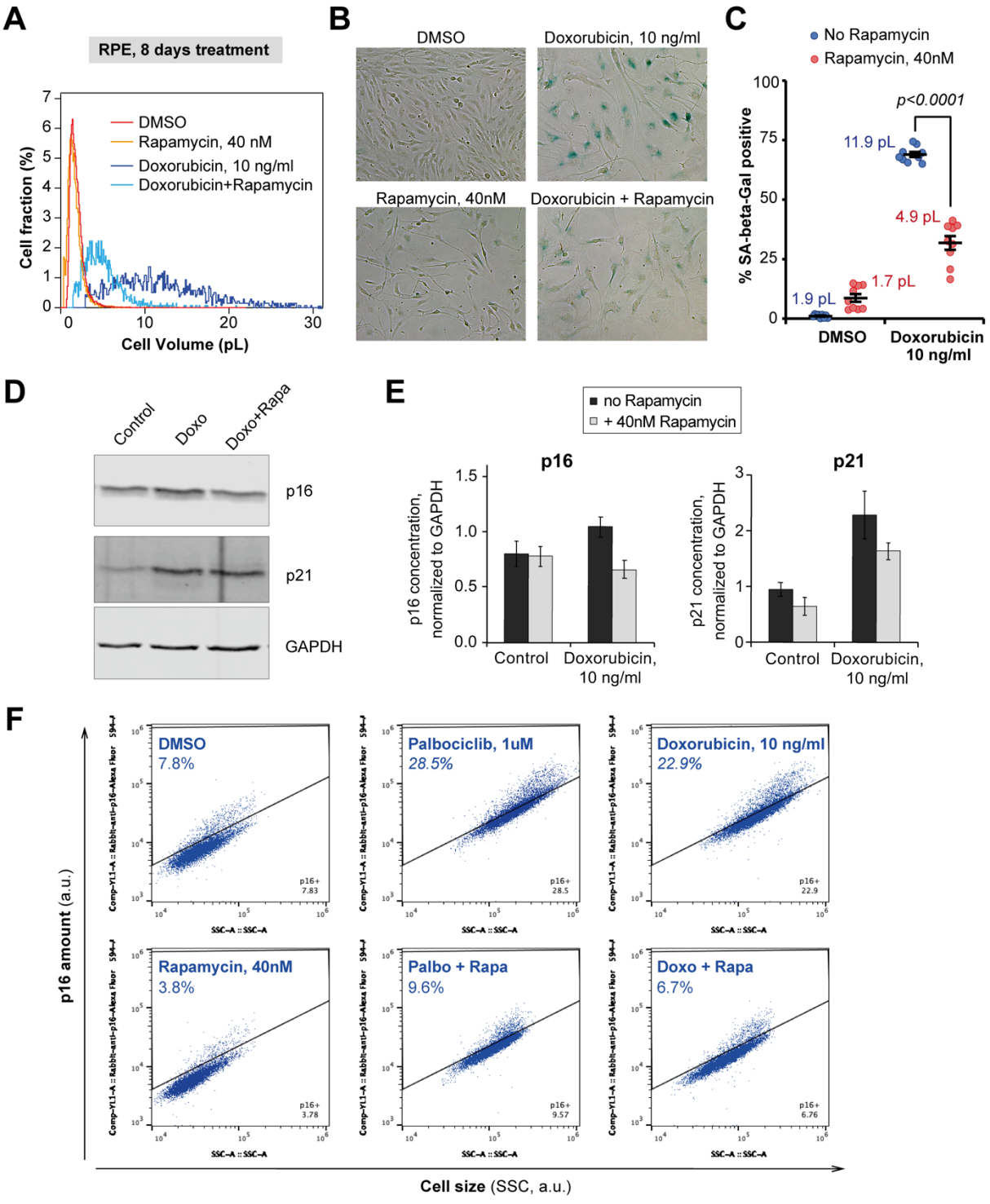

Figure S14. Anti-senescent effects of Rapamycin in Doxorubicin-treated cells, related to Figure 5: (A, B) Cell size distributions (A) and characteristic images showing SA-beta-Gal staining (B) of RPE-1 cells treated for 8 days with DMSO or 10 ng/ml Doxorubicin in the presence

or absence of 40nM Rapamycin. **(C)** Effect of Rapamycin, which reduces cell growth, on the percentage (±standard error) of SA-beta-Gal positive cells in RPE-1 cultures treated with Doxorubicin for 8 days. SA-beta-Gal quantification for every data point included 700-1200 cells quantified from 9 different fields of view. Values shown next to each condition indicate the mean cell sizes after 8 days of treatment. **(D-E)** Effects of Rapamycin on the expression of senescence markers p21 and p16 in HLF cells that were treated with Doxorubicin for 8 days. (D) A representative immunoblot against p21 and p16. (E) Quantification of p21 and p16 immunoblots. Data are shown as mean ± standard error, n = 4 biological replicates. **(F)** Percentage of HLF cells expressing high levels of p16 after 8 days of Doxorubicin (10 ng/ml) or Palbociclib (1μM) treatment in the presence or absence of Rapamycin (40 nM). Measurements were made using immunofluorescent staining followed by flow cytometry.

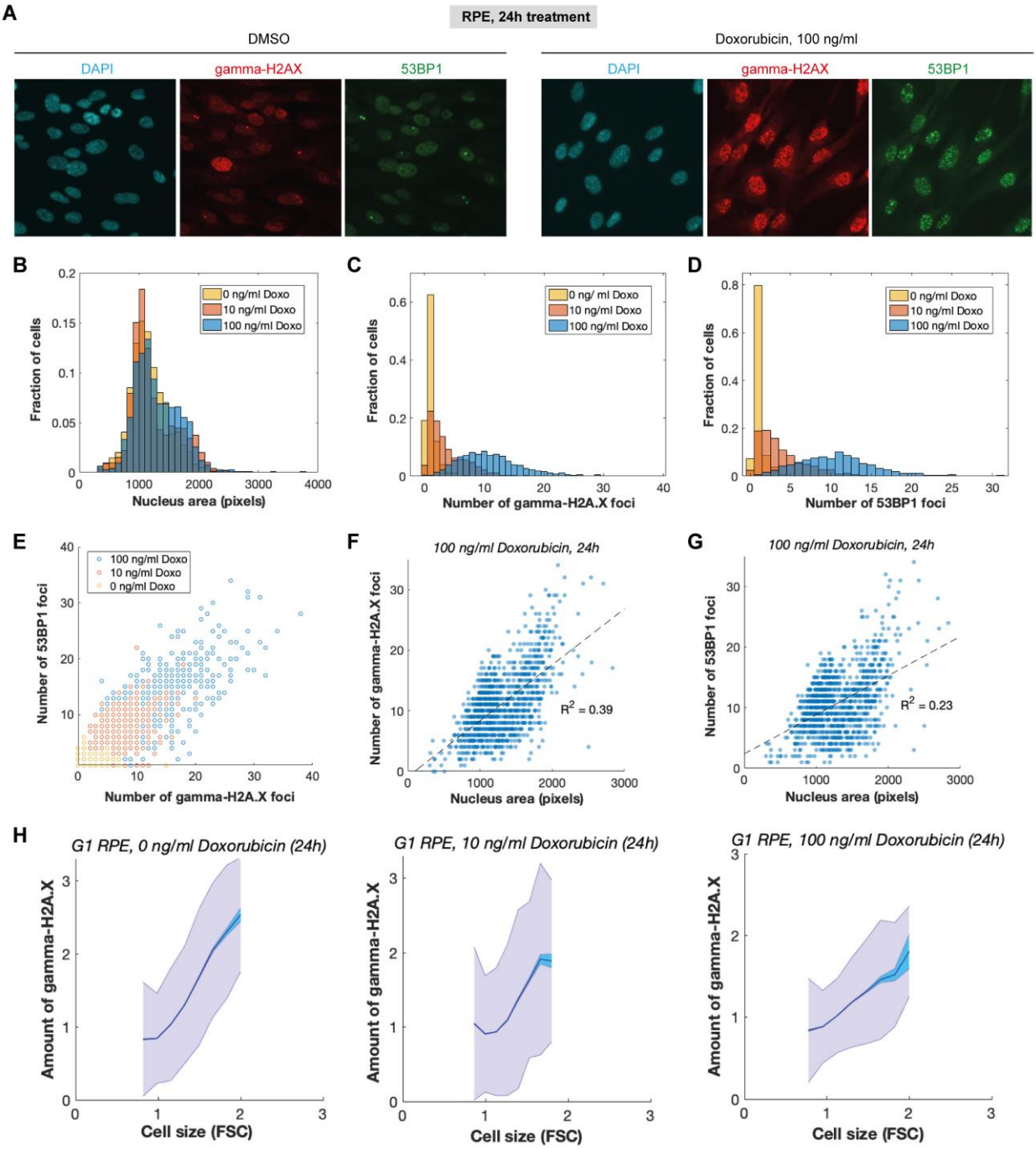

Figure S15. Correlation between cell size and number of DNA damage foci in RPE-1 cells, related to Figure 6. (A) Immunofluorescent staining of RPE-1 cells treated with DMSO or 100 ng/ml Doxorubicin for 24 hours against γ-H2AX (red) and 53BP1 (green), with DAPI staining shown in cyan. **(B-D)** Distribution of nuclear areas (B) and numbers of γ-H2AX (C) and 53BP1

(D) foci per nucleus in RPE-1 cells treated with DMSO, 10 ng/ml or 100 ng/ml Doxorubicin for 24 hours. **(E)** Correlation between the γ-H2AX and 53BP1 foci numbers in each RPE-1 nucleus. **(F-G)** Correlation between the nuclear area and the number of γ-H2AX (F) and 53BP1 (G) foci per nucleus in each RPE-1 cell nucleus after treatment with 100 ng/ml Doxorubicin for 24 hours. n = 1602 cells for DMSO, n = 1265 cells for 10 ng/ml Doxorubicin, and n = 1062 cells for 100 ng/ml Doxorubicin. (**H**) Amounts of γ-H2AX in G1-phase RPE-1 cells, measured using flow cytometry and plotted against cell size (FSC). The cells were treated with DMSO (left), 10 ng/ml Doxorubicin (center), or 100 ng/ml Doxorubicin (right) for 24 hours prior to cell staining. The antibody staining and FSC signals were normalized to mean values for each individual sample. The data were binned by cell size (FCS), and bin means (solid blue lines), standard errors of the mean (light-blue shaded area), and standard deviations (purple shaded area) were plotted against cell size. n = 100,000 cells were analyzed for each condition.