bioRxiv preprint doi: https://doi.org/10.1101/2021.05.03.442497; this version posted May 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Revisiting the Hayflick Limit: Insights from an Integrated Analysis of Changing Transcripts, Proteins, Metabolites and Chromatin.

Michelle Chan, Han Yuan*, Ilya Soifer*, Tobias M. Maile, Rebecca Y. Wang, Andrea Ireland, Jonathon O'Brien, Jérôme Goudeau, Leanne Chan, Twaritha Vijay, Adam Freund, Cynthia Kenyon, Bryson Bennett, Fiona McAllister, David R. Kelley, Margaret Roy, Robert L. Cohen, Arthur D. Levinson, David Botstein**, David G. Hendrickson**
*contributed equally
**Corresponding Authors

Abstract

Replicative senescence (RS) as a model has become the central focus of research into cellular aging *in vitro*. Despite decades of study, this process through which cells cease dividing is not fully understood in culture, and even much less so *in vivo* during development and with aging. Here, we revisit Hayflick's original observation of RS in WI-38 human fetal lung fibroblasts equipped with a battery of high dimensional modern techniques and analytical methods to deeply profile the process of RS across each aspect of the central dogma and beyond. We applied and integrated RNA-seq, proteomics, metabolomics, and ATAC-seq to a high resolution RS time course. We found that the transcriptional changes that underlie RS manifest early, gradually increase, and correspond to a concomitant global increase in accessibility in nucleolar and lamin associated domains. During RS WI-38 fibroblast gene expression patterns acquire a striking resemblance to those of myofibroblasts in a process similar to the epithelial to mesenchymal transition (EMT). This observation is supported at the transcriptional, proteomic, and metabolomic levels of cellular biology. In addition, we provide evidence suggesting that this conversion is regulated by the transcription factors YAP1/TEAD1 and the signaling molecule TGF-β2.

1 Introduction

Replicative senescence (RS) of animal cells growing *in vitro* was first discovered by Leonard Hayflick. He found that primary human diploid fibroblast cell lines ceased to proliferate after an extended number of serial passages [1]. Since then, considerable work has been done to describe this phenomenon. A major causal feature of RS is telomere erosion, a process in which the telomeres of the chromosomes gradually shorten with increasing cellular divisions eventually triggering the DNA damage response and cell cycle exit [2]. This is understood to be due the absence of the telomerase reverse transcriptase (hTERT), the catalytic component of human telomerase which adds telomeric sequences to the ends of chromosomes to maintain telomere length in germ cells and stem cells. As hTERT activity is undetectable in normal human somatic cells, telomere attrition is a common aging phenotype hypothesized to underlie cellular senescence at the organismal level [3, 4]. Replicative senescence of human somatic cell lines *in vitro* can be avoided by overexpression of hTERT which prevents telomere shortening and confers apparently unlimited replicative capacity [5].

Beyond growth arrest and telomere shortening, phenotypic changes exhibited in replicatively senescent cells include the senescence associated secretory phenotype (SASP). The SASP proteins include proinflammatory cytokines, growth factors, angiogenic factors, and proteases, all of which are thought to play a role in paracrine signaling whereby senescent cells can promote local wound healing and/or drive healthy neighboring cells into senescence [6, 7, 8] [9]. Replicatively senescent cells also accumulate DNA and protein damage, accumulate lipids, and lose regulatory control of mitochondria and lysosomes [10].

The phenotypic similarity (telomere attrition, epigenetic alterations, mitochondrial dysfunction and loss of proteostasis) between the cell autonomous aging hallmarks and *in vitro* senescence has led to the hypothesis that senescent cells *in vivo* play a causal role in organismal aging and aging-related diseases [11, 12, 13, 14, 15, 16, 3].

For example, up to 15% of familial idiopathic pulmonary fibrosis cases arise from mutations in telomerase. In addition, up to 25% of sporadic cases have telomere lengths less than the 10th percentile. These findings point to a

bioRxiv preprint doi: https://doi.org/10.1101/2021.05.03.442497; this version posted May 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

potential role for replicative senescence in IPF pathology. [17, 18, 19, 20, 21, 22, 23, 24, 25]. Furthermore, the elimination of senescent cells in a number of age-related diseases, such as cardiac fibrosis, pulmonary fibrosis, neurodegenerative diseases, osteoporosis, and metabolic disorders have been argued to alleviate the disease state [26]. Clinical trials for senolytics targeting fibrotic diseases, osteoporosis, frailty, and metabolic syndromes are currently underway [27].

The mechanism by which senescent cells might contribute to aging phenotypes is currently still unclear [11, 28]. In an effort to bring clarity to RS *in vitro* that in turn could elucidate *in vivo* function, we revisited and redesigned the original Hayflick experiment. Making use of recent advances in high-dimensional technologies, (bulk RNA-seq, single cell RNA-seq (scRNA-seq), ATAC-seq, metabolomics and proteomics) we tracked changes throughout the replicative lifespan of the original Hayflick WI-38 cell line.

Overall, our data recapitulate many known features of the *in vitro* senescence process while simultaneously providing novel insight. First, the time resolution of our experiment coupled with single cell trajectory analysis reveals that senescence is a gradual process that shares transcriptional, proteomic, and metabolomic features with epithelial-mesenchymal transition (EMT). Second, our metabolomic data identifies Nicotinamide N-methyltransferase (NNMT) activity as a potential initiating event in RS dependent heterochromatin loss. Third, we show that the increased accessibility in heterochromatic regions is concentrated in nucleolar/lamin associated domains and corresponds with observed changes in the RS transcriptome. Lastly, integration across data modalities reveals that RS replicatively senescent WI-38 cells bear a strong resemblance to myofibroblasts. We provide evidence that the YAP1/TEAD1 transcription factor complex and TGF-β2 signaling are putative regulators of the transition to this state. Together our data suggests that a process similar to fibroblast to myofibroblast transition (FMT; analogous to EMT) underlies is an intrinsic aspect of the RS phenotype in WI-38 fibroblasts.

2 Results

2.1 Transcriptomic profiles of replicative senescence (RS), radiation induced senescence (RIS) and increasing cellular density (CD)

To capture the replicative senescence process with high resolution we designed an experiment to continuously grow and intermittently sample cells from a starter batch of WI-38 cells that had undergone only 20 population doublings (PDL 20). To distinguish between RS-dependent changes and those arising from altered cell density and growth rate we performed a cell density control study (CD). Briefly, early PDL cells were sampled at increasing levels of cell density to measure gene expression changes associated with cellular density and decreasing cell proliferation independent of replicative senescence (Figure 1A). In addition to the CD control, we also included TERT immortalized WI-38 cells (hTERT) grown in parallel as a control for long term culturing and day-to-day sampling batch effects [5]. We also included a control for replication-independent senescence using radiation induced senescence (RIS) to test for differences between RS and RIS. Finally, we sampled proliferating WI-38 cells at multiple PDLs for RNA-seq, scRNA-seq, proteomics, metabolomics and ATAC-seq until the WI-38 cells had reached senescence as measured by the cessation of growth (Figure 1B). As expected, hTERT immortalized cells grew at a constant rate and did not slow or cease growth (Figure 1B).

We first examined the bulk RNA-seq data to compare and contrast RS with RIS, CD, and hTERT cells. Differential gene expression analysis using RNA-seq revealed 8,968 genes change with increasing PDL, 8,690 genes change in response to ionizing radiation, and 11,652 genes vary with increasing cell density (SData 1). Notably, the RS phenotypes and transcriptional changes we observed were consistent with previous studies. Figure 1C shows the strong correlation of the pattern of induced transcription established in PDL 50 cultures and the fibroblast derived senescence-associated signature reported by Hernandez-Segura et al(2017) [28]. We also observed induction of senescence-associated-β galactosidase (Figure 1D) as well as p16 and p21 (Figure S1).

Hierarchical clustering of all significantly changing genes across the four conditions highlighted several important features of our experiment. First, as expected, gene expression in the hTERT-immortalized WI-38 cells remained largely stable. Second, RS, RIS, and CD exhibited many shared, but also unique, gene expression changes with respect to both the identity of differentially expressed genes and the magnitude of their changes (Figure 1E).

To facilitate a biological interpretation of these shifting transcriptomic landscapes, we applied Gene Set Enrichment Analysis (GSEA) analysis using the MSigDB Hallmark annotation sets to learn which general processes are shared and distinct between RS, RIS, and CD (Figure 1F, SData 2) [29, 30, 31]. All three perturbations (and not the immortalized cells), exhibited dramatic reductions in expression of genes belonging to S and M cell cycle phases, consistent with cessation of cell division (Figure 1F - top cluster). Perhaps driven by this shift, we also found significant overlap in additional enriched annotations among genes induced in RS, RIS and CD, albeit with some variation in significance. Many of the shared enrichments can be categorized as stress responses, e.g. apoptosis, p53 pathway, inflammatory and interferon responses, and STAT3/5 signaling. In addition, we also observed enrichment of several development/differentiation gene sets, including myogenesis, angiogenesis, and adipogenesis.

We observed relatively few gene sets with discordant patterns across RS, RIS, and CD (Figure 1E). However, the gene set for Epithelial to Mesenchymal Transition (EMT) was significantly enriched in genes that increased with RS as opposed to RIS ($p = 5.5e-6$ vs $p=0.12$) or CD wherein the term was actually underrepresented, but not significantly so. Although RS, RIS and CD are highly similar at the abstracted level of enriched gene sets, it is clear in (Figure 1E) that there are many induced genes specific to RS. In addition, we applied GSEA to each individual time point for each condition (Figure S2,SData 2) and found the EMT gene set is enriched early and robustly during

bioRxiv preprint doi: https://doi.org/10.1101/2021.05.03.442497; this version posted May 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

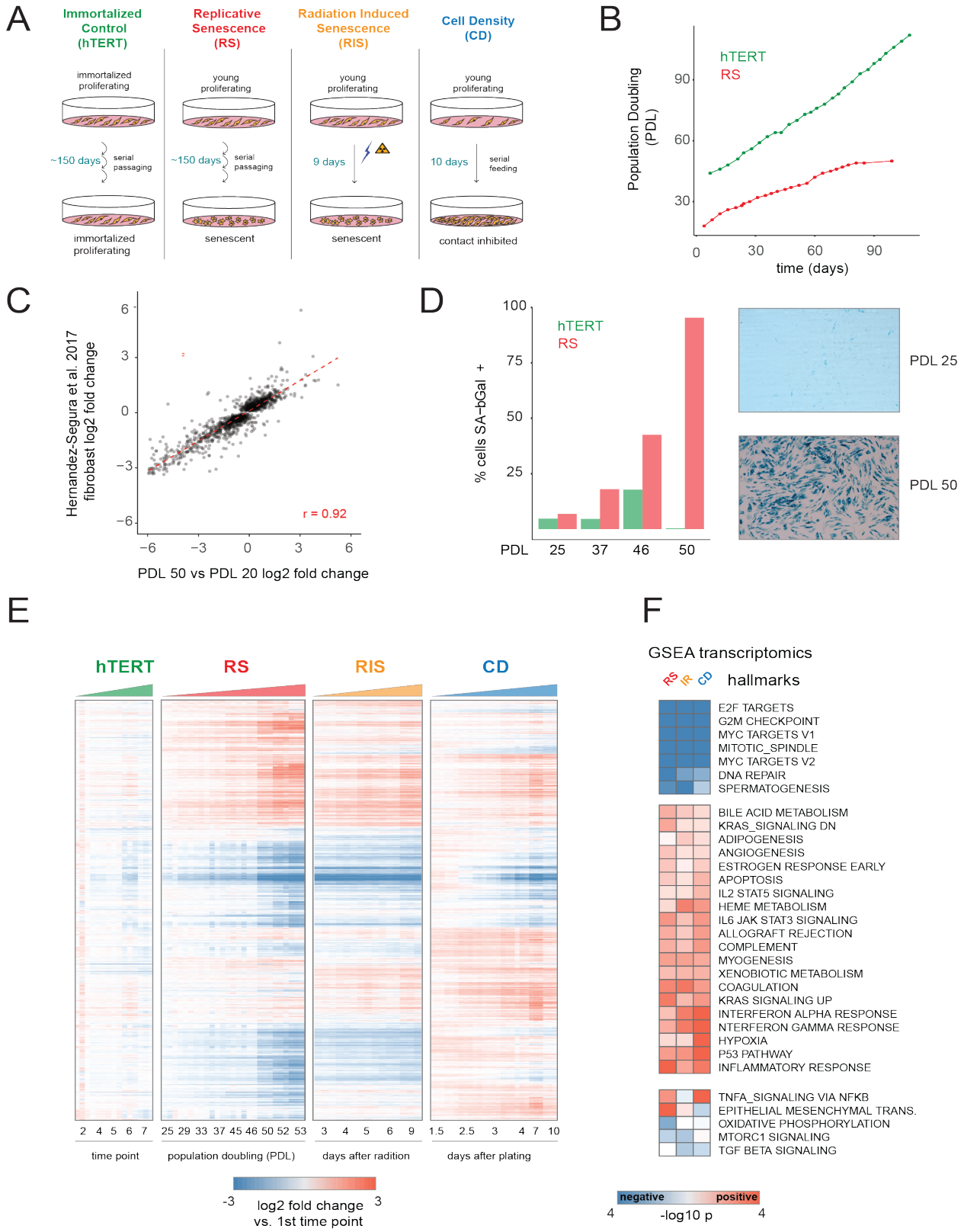

Figure 1: **Expression dynamics of replicative senescence (RS), radiation induced senescence (RIS), increasing cell density (CD) and hTERT WI-38 cells. (A)** Experimental design for the RS, RIS, CD and hTERT experiments. **(B)** Scatterplot of days in culture (x-axis) vs. population doublings (PDL)(y-axis) for WT WI-38 cells (red) and hTERT immortalized cells (green). **(C)** Scatterplot of gene expression changes ($\log 2$(fold change)) for PDL 50 vs PDL 20 (x-axis) vs. senescence log2 fold change derived from a generalized linear model (GLM) compiling several fibroblast cell lines (Hernandez-Segura et.al. 2017). **(D)** (right) Percent cells staining positive for SA-βgal staining (y-axis) at increasing PDLs (x-axis) for WI-38 cells (red) and hTERT (green);(left) **(E)** Hierarchical clustering of significant genes expression changes across all conditions. Values are log2 fold change vs. the average of the first time point of each condition. **(F)** Gene Set Enrichment Analysis (GSEA) results for RS, RIS and CD using the MsigDB Hallmarks annotation set. The -log10 p-value is colored by direction of enrichment (red=up, blue=down).

bioRxiv preprint doi: https://doi.org/10.1101/2021.05.03.442497; this version posted May 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

RS. Thus the EMT gene set appears to represent a particularly important aspect of RS biology.

2.2 Single-cell RNA-seq reveals that replicative senescence is a gradual process

Interestingly, the vast majority of gene expression changes evident by PDL 50 begin to manifest at much earlier PDLs. This observation is consistent with two distinct possibilities: (A) The senescence expression changes accrue early and gradually in the majority of cells without respect to proliferation status, or (B) the bulk RNA-seq profiles are a changing admixture of transcriptionally distinct mitotic, G1, and senescent cells. Both models have substantial support in the literature. [32, 33, 34, 35, 36] To discriminate between these possibilities, we employed single-cell RNA-seq (scRNA-seq) to directly measure percentages senescent and cycling cells.

Briefly, we collected between 1,000-2,000 cells at increasing PDLs of WI-38s and matched hTERT time points (SData 3). We applied a modified (see methods) 10X Genomics protocol to capture and preserve the more fragile senescent cells (Figure S3). We profiled a total of 15,000 cells (9,000 WT and 6,000 hTERT). When aggregated into pseudo-bulk profiles (i.e the sum the single-cell reads) expression profiles, the single cell results are highly correlated with expression changes observed in bulk RNA-seq (r=0.77-0.83,Figure S4).

We first classified cells as either S, G2M, or G1 phase using canonical markers for the S and G2M phases as previously described [37]. As expected, based on bulk transcriptomics, the number of cells in either S or G2M decreases with increasing PDL whereas the cycling cell proportions are stable in the hTERT timepoints (Figure 2A,B, SData 3).

Next, we projected all wild type (WT) cells and two hTERT time points together in one UMAP to identify broad patterns in replicative aging PDL cells versus cycling hTERT cells. Cells from the two cell lines organize separately, however share a similar geometry composed of a S/G2M phase roundabout and a large G1 lobe (Figure 2C). Although the overall pattern of WI-38 cells for both cultures was highly similar, we observed progression of cell grouping with PDL that had no concordance with increasing hTERT time points (Figure 2D)- red to blue). Specifically, PDL50 cells (but not the hTERT time point control) cluster apart from all other cells. We identified this cluster as "senescent" by scoring each cell with our senescent gene expression signature derived from our bulk RNA-seq (Figure S5, SData 3).

To test model A vs B, we compared the pseudo-bulk gene expression profiles across the cell cycle at each PDL. We reasoned that if RS were gradual, we would observe the RS gene expression pattern even in young cycling cells without respect to G1/S/G2M. Indeed, the RS signature is present in all cell cycle phases even in early PDLs (Figure 2E- left panels vs. far right panel). In addition, isolation and UMAP projection of cells from each cell cycle phase revealed that the dominant source of variance in cycling cells is the PDL (Figure S6). Lastly, we did not observe any cells < PDL50 in close association with the senescent cells in the UMAP projection PDLs (Figure 2C,D). Together, these results argue in favor of a gradual model (A) of replicative senescence wherein cells ramp up expression of the RS program with increasing PDL even when still proliferative. These data are consistent with the previously reported changes in cellular phenotype (larger cell size and increased cycling time) with increasing PDL. [38, 39, 40, 41]. Importantly, this observation suggests that aspects of cellular senescence are present in non-senescent cells. This result raises the intriguing possibility that the reported pathological features of senescent or senescent-like cells *in vivo* might also manifest in cells that are not classically senescent.

2.3 Proteomic landscape of RS WI-38 cells suggests large metabolic alterations

We next turned to our proteomics data generated from the same cultures and timepoints as above. We obtained high confidence measurements for 8500 proteins (SData 4). Similar to our transcriptional results, the TERT samples did not exhibit large changes in proteome with experimental progression (Figure S7). Overall, we observed high concordance between transcript and protein levels with very few outliers (r=0.71 for PDL50 vs PDL20, Figure S8). We employed GSEA using hallmark annotations and observed a sharp depletion of proliferation and mitosis associated gene sets supporting our previous findings (Figure 3A, SData 4). Likewise, the proteomics data recapitulated enrichment of the EMT annotation set along with the adipogenesis and myogenesis.

Multiple enriched sets pointed to RS-dependent shifts in cellular energy utilization. (Figure 3A). Of note, we found that although under-enriched in our RS transcriptome (Figure 1F), The oxidative phosphorylation hallmark was highly enriched in the proteomics data. We plotted the oxidative phosphorylation genes driving the set enrichment in Figure S8. This projection revealed that many genes in this set fall into a quadrant (positive proteomic, negative transcriptomic) suggesting discordant regulation of these genes.

The oxidative phosphorylation hallmark set is composed of multiple closely linked mitochondrial complexes and functions that regulate cellular energy flux e.g. TCA cycle, fatty acid oxidation, pyruvate metabolism, ATP synthase etc. We divided the oxidative phosphorylation gene set into into these constituent parts and visualized the changes in proteins with increasing PDL. We observed increases in all categories except for mitochondrial assembly and structures (Figure S9. SData 5).

Next, we examined the annotations for the highest enriched hallmark sets from Figure 3A and found strong enrichment of multiple KEGG metabolic pathways. (Figure 3B). In addition, both the glucose transporter SLC2A1/GLUT1 and the fatty acid scavenger CD36 transcripts exhibited strong, early, and RS specific up-regulation (Figure 3C). Together these results point to a drastically altered RS metabolic landscape.

bioRxiv preprint doi: https://doi.org/10.1101/2021.05.03.442497; this version posted May 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

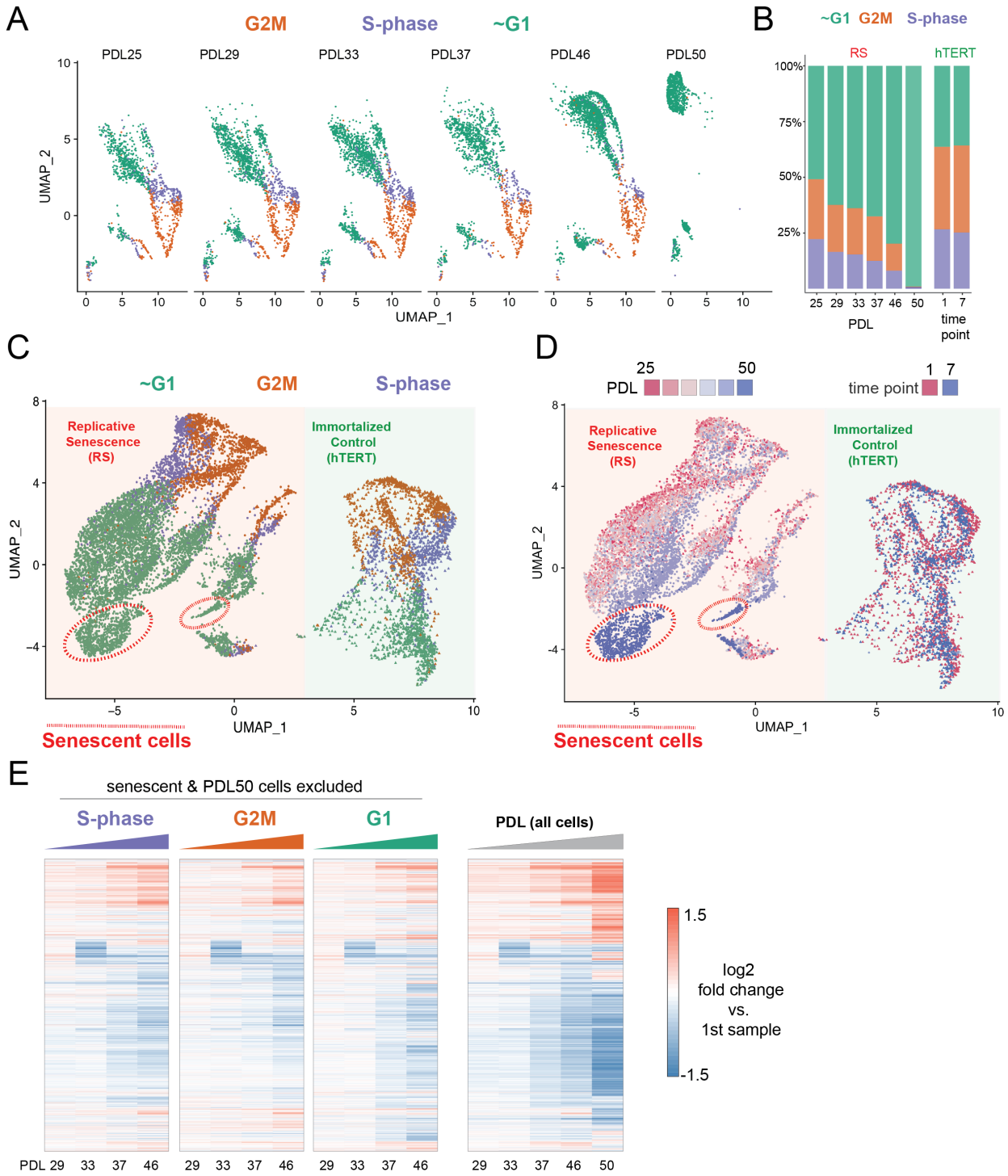

Figure 2: **Cell cycle exit and distribution on approach to replicative senescence does not explain gradual increase in the RS transcriptome.** **(A)** Individual UMAP projections of WT WI-38 cells by PDL colored by phases of the cell cycle (G1 = green, G2/M = orange, S-phase = purple). **(B)** Bar graph of cell cycle state percentages (y-axis) by PDL (x-axis) for WT WI-38 cells (left) and hTERT WI-38 cells (right). **(C)** UMAP projections of WT and hTERT WI-38 cells by PDL colored by different phases of the cell cycle. **(D)** UMAP projection of all WI-38 cells from increasing PDLs (RS) or time point (hTERT)-colored from early (red) to late (blue). **(E)** The RS transcriptome manifests early in all phases of cell cycle. Heatmap of hierarchical clustering of gene expression values of differentially expressed genes as aggregated transcriptomic profiles for each cell cycle phase and PDL (left) versus all cells (right). Values are log2 fold change of each PDL against the first (not shown).

bioRxiv preprint doi: https://doi.org/10.1101/2021.05.03.442497; this version posted May 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

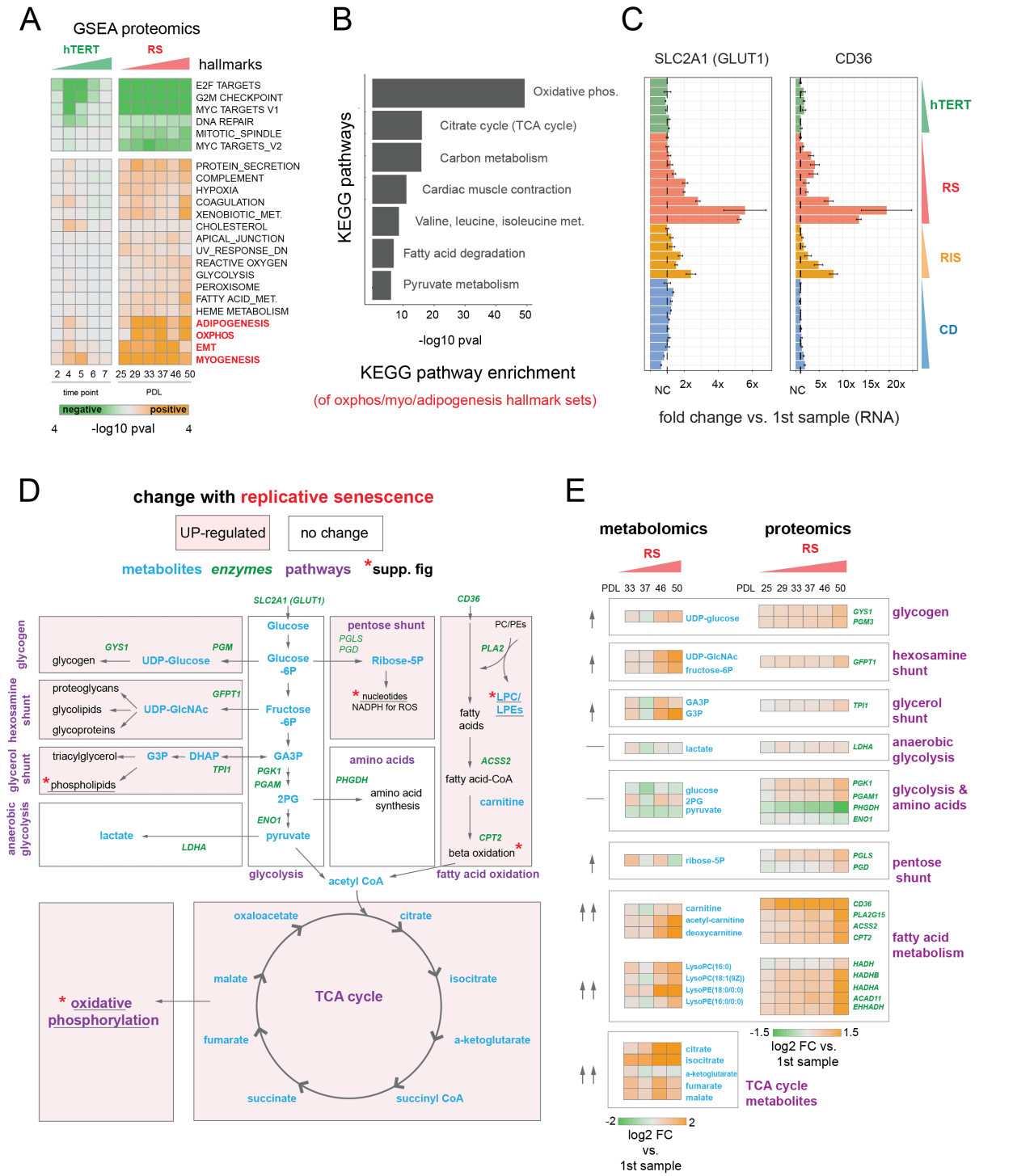

Figure 3: **Proteomic and metabolomic changes during replicative senescence reveal increased utilization of oxidative phosphorylation, glycolytic shunts, and fatty acid oxidation (A)** Significant ($p < 0.01$) Gene Set Enrichment Analysis (GSEA) results for protein expression changes using the MsigDB Hallmarks annotation set. The -log10 p-value is plotted and colored by the directionality of enrichment (orange=up, green=down). **(B)** Kegg pathway GO term enrichment of leading edge genes from A for Oxidative Phosphorylation, Myogenesis and Adipogenesis. **(C)** Gene expression fold changes (replicate average) for the insulin-independent glucose transporter SLC2A1 (left) and fatty acid transporter CD36 (right) in the hTERT, RS, RIS and CD time courses. **(D)** Metabolic diagram of glycolysis, glycolytic shunts, fatty acid import and oxidation, TCA cycle/Oxidative Phosphorylation. Metabolites are blue, proteins are italicized in green. * denotes metabolite classes expanded in supplementary figures. Shading indicates inferred direction of pathway during RS based on metabolite/protein changes in E. **(E)** Heatmaps of metabolite and protein changes with RS from D (replicate averages).

bioRxiv preprint doi: https://doi.org/10.1101/2021.05.03.442497; this version posted May 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

2.4 RS WI-38 cells exhibit profoundly increased utilization of fatty acid metabolism and glycolytic shunts

To generate a metabolic profile for RS in WI-38 cells, we harvested samples at increasing PDLs alongside input material for all other data types (methods, SData 6). Out of 288 detected and identified compounds, we focused on compounds exhibiting PDL-dependent changes in abundance > log2 0.5. To guide further our metabolic analysis, we focused our queries based on the enriched metabolic pathways found in the proteomic analysis. Specifically we studied the changes to glycolysis, oxidative phosphorylation and fatty acid metabolism (Figure 3D).

Following the glucose, we looked at the central stem of glycolysis as a potential source of energy supplying the observed increases in the TCA cycle and oxidative phosphorylation (Figure 3D,E). We found that metabolites and enzymes dedicated to pyruvate generation and/or lactate production did not change concordantly. We instead found that most of the glycolytic shunts exhibited up-regulation at both the metabolite and protein level (Figure 3E). These results suggest an increase in allocation of glucose for manufacture of various biomolecules and their precursors (glycogen, hexosamines, phospholipids) (Figure 3E and Figure S10).

In contrast, metabolites and enzymes involved in fatty acid import and oxidation for the purpose of energy generation appear strongly up-regulated (Figure 3E). We also saw that metabolites involved in the *de novo* production of phospholipids via the Kennedy pathway are highly up-regulated, specifically phosphotidylethanolamines precursors (Figure S11).

From these data, it is clear that RS WI-38 cells undergo drastic shifts in metabolism. Specifically, we see increased glucose utilization in glycolytic shunts coupled with an increase in fatty acid import and oxidation. It is possible that RS cells are switching to fatty acid oxidation to fuel increased TCA cycling and oxidative phosphorylation as glucose is diverted to macromolecule production. Indeed, metabolomic data collected from various types of senescence models has shown increased lipid oxidation, lipid accumulation, TCA up-regulation, and glycolytic alterations [42, 43, 44, 45, 46]. Likewise, consensus metabolomic findings in EMT models report increased TCA cycle products, altered lipid metabolism, and activated hexosamine pathway [47, 48, 49]. These data provide functional metabolic evidence supporting a connection between RS and the EMT hallmark enrichment observed in RNA and protein expression.

2.5 Nicotinamide N-methyltransferase (NNMT) links nicotinamide adenine dinucleotide (NAD) and methionine metabolism as a putative heterochromatin regulator

Metabolic regulation of epigenetic state is an increasingly recognized mechanism through which nutrient availability influences cellular function by regulating the abundance of cofactors required for histone modifications [50]. Specifically, Nicotinamide adenine dinucleotide (NAD+) and methionine metabolism power the deacetylation and methylation required for maintaining repressive DNA conformations [51]. In a compelling intersection, perturbation of NAD, methionine, and heterochromatin levels have all previously been reported in multiple aging contexts including RS [52, 53, 54, 55, 56]. In compiling our RS data for the metabolic and proteomic components of the NAD and methionine pathways, we found that one of the largest and earliest changes for any protein (or transcript) was the increased expression of Nicotinamide N-methyltransferase (NNMT) (Figure 4A,B)).

DNA and histone methylation require abundant levels of the universal methyl donor S-adenosyl methionine (SAM). NNMT not only depletes SAM by catalyzing the removal of the SAM methyl group, it does so by fusing it to the NAD+ precursor nicotinamide (NAM) resulting in the production of the inert molecule methyl nicotinamide (MNA) [57]. Thus, NNMT effectively acts as a sink for the two primary metabolites a cell requires for silencing gene expression and forming heterochromatin [58]. We found that MNA levels mirror those of NNMT; increasing early and robustly during RS progression consistent with high NNMT activity. SAM and its sans-methyl version (S-adenosyl homocysteine (SAH)) were depleted with RS albeit to a lesser extent than the observed increase in MNA. NAMPT and NMNAT1, enzymes in the NAD pathway, were weakly down-regulated at the protein level (Figure 4A,B)).

High NNMT expression has been implicated in methyl depletion in embryonic stem cells and Cancer-Associated fibroblasts (CAFs) [59, 60]. In both cases the functional consequences were similar; DNA hypomethylation and decreased capacity to form or maintain heterochromatin. In both examples, the important point is that the loss of silencing is actively promoted through NNMT activity. One intriguing hypothesis is that increased NNMT activity with RS directly induces heterochromatin loss via SAM depletion (Figure 4C).

2.6 Increased DNA accessibility and transcription from nucleolar and lamin associated domains is a dominant feature of the RS epigenome

Previous reports have described large scale senescence-induced alterations to the epigenome including but not limited to heterochromatin loss, formation of heterochromatin foci (SAHF) and perturbation of Nucleolar-Associated Domains (NADs) and Lamin-Associated Domains (LADs) [61, 62, 63]. With the advent of ATAC-seq, a transposition based method for the measurement of DNA accessibility, it is now possible to study these phenomena with relative ease across multiple timepoints [64]. Furthermore, recent analytical advances have facilitated the interrogation of epigenetic state via the distillation of multidimensional epigenetic data into discrete chromatin states (e.g. heterochromatin, promoter, enhancer etc.) across hundreds of cell lines [65].

To study genome-wide changes in the epigenetic landscape of senescent cells, we collected ATAC-seq data during the RS time course and quantified the read distribution across chromatin states previously annotated for the IMR-90 human fetal lung fibroblast cell line [65]. We categorized 25 distinct states into 4 broad categories: promoters, enhancers, transcription, and mis-

bioRxiv preprint doi: https://doi.org/10.1101/2021.05.03.442497; this version posted May 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

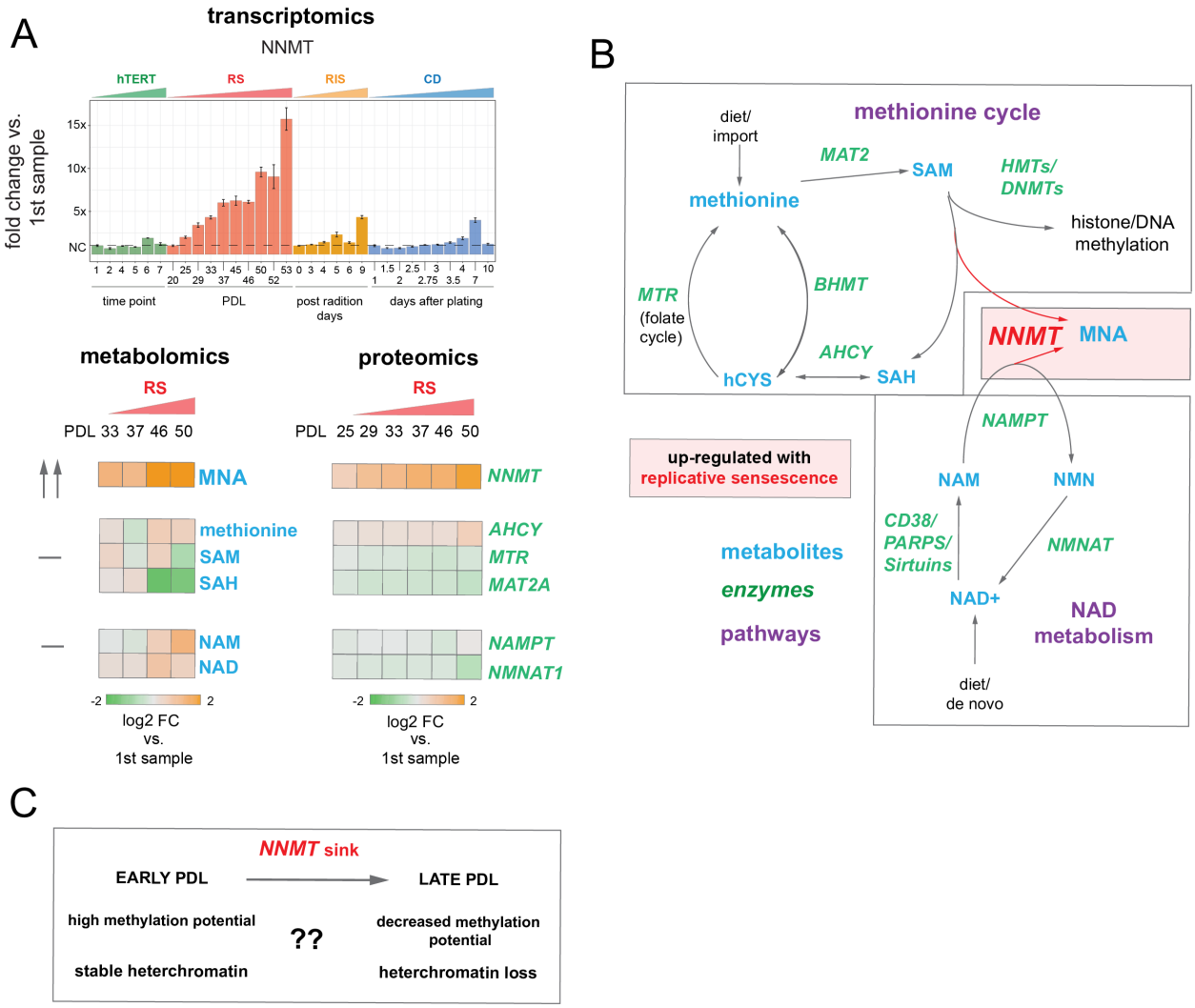

Figure 4: **Nicotinamide N-methyltransferase (NNMT) Links Nicotinamide adenine dinucleotide (NAD) and Methionine Metabolism in RS (A)** Gene expression fold changes for the NNMT (top) in the hTERT, RS, RIS and CD time courses (replicate average). Heatmaps of metabolite and protein changes during RS from B (bottom-replicate average). **(B)** Metabolic diagram of the methionine and NAD salvage pathways. Metabolites are blue, proteins are italicized in green. Shading indicates inferred pathway direction during RS based on metabolite/protein changes in A. **(C)** Potential model for NNMT and metabolic regulation of heterochromatin during RS.

bioRxiv preprint doi: https://doi.org/10.1101/2021.05.03.442497; this version posted May 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

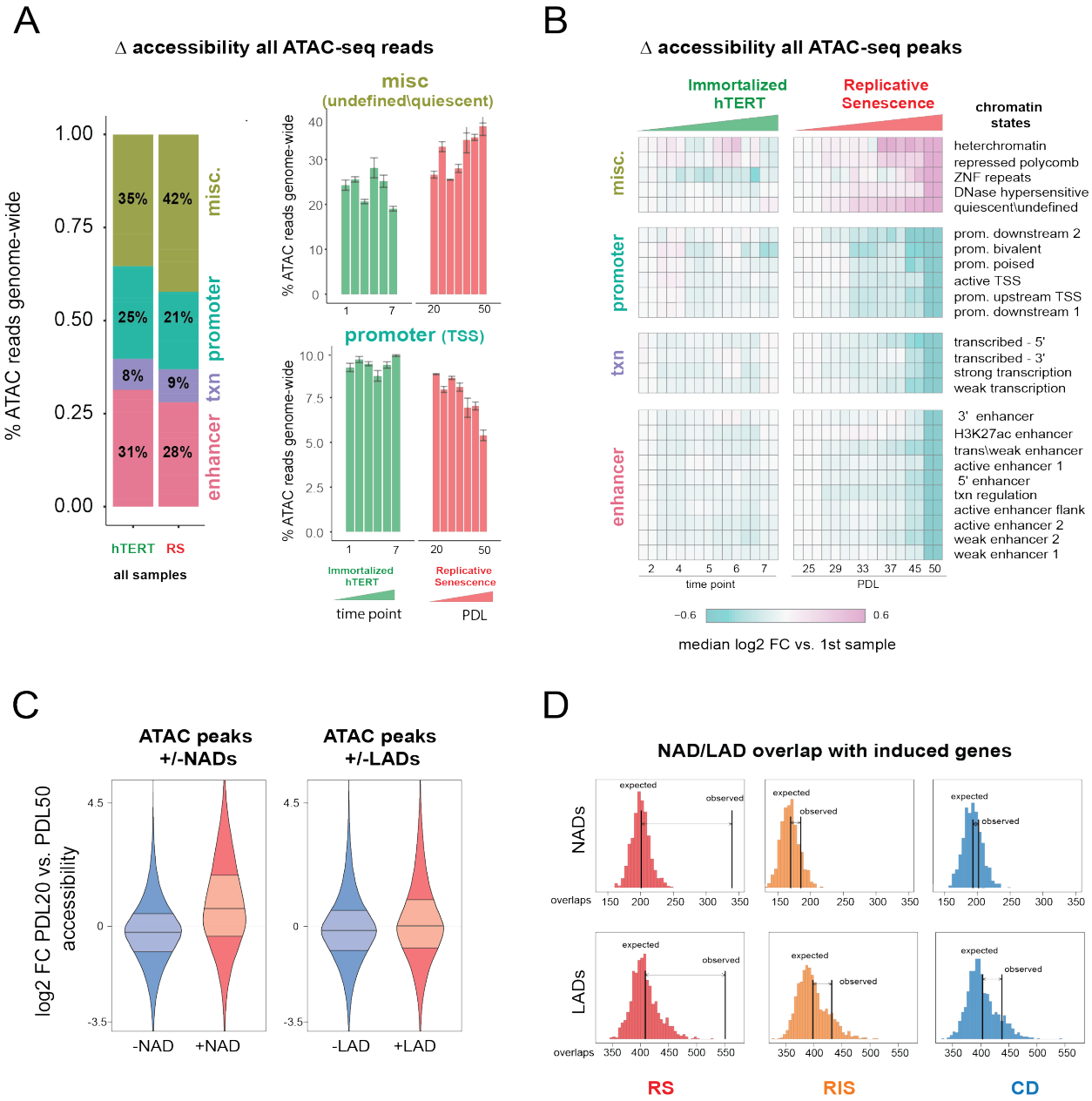

Figure 5: **Increased Accessibility within heterochromatin and Nucleolar Associated Domains (NADs) is a dominant feature of the RS epigenome (A)** Bar graph of percent of ATAC-seq reads falling into 4 broad chromatin states compiled from the ENCODE IMR-90 25 chromatin state prediction for all samples in WT or hTERT WI-38 cells. Bar graphs displaying percent of all ATAC-seq reads within two specific states (y-axis) vs PDL (x-axis) for WT (red) and hTERT (green) WI-38 cells. Error bars represent standard deviation. **(B)** (left) Log2 fold change in counts vs. first time point for significantly changing ATAC-seq peaks binned into ENCODE IMR-90 25 chromatin states. **(C)** Log2 fold change distribution (PDL 50 vs PDL 20) for all ATAC-seq peaks that overlap (pink) or are separate from (blue) from NADS (left) or LADs (right). **(D)** Randomized permutation test distributions for highly induced genes (from Figure 1) in RS (red), RIS (orange) and CD (blue). Median expected number of overlaps and observed number of overlaps shown by black bars. (top) for NADs, (bottom) for LADs.

bioRxiv preprint doi: https://doi.org/10.1101/2021.05.03.442497; this version posted May 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

cellaneous, the last of which is composed largely of heterochromatic and undefined states, e.g. H3K9me3 and H3K27me3 rich regions, zinc-finger repeats, and quiescent–defined by the absence of histone marks, accessibility, or gene features [65].

In comparing the read distribution into chromatin states between hTERT immortalized cells and WT WI-38 cells of increasing PDL, we observed an increase in ATAC-seq reads falling into the miscellaneous chromatin regions at the relative cost to all other states (Figure 5A-left panel). We plotted the read distribution across chromatin states as a function of increasing time for both cell lines and found these shifts to be associated with increasing PDL in the WT WI-38 cells; hTERT ATAC-seq read proportions across chromatin states remained stable over time (Figure 5A-right panel).

To determine if this shift in accessibility to undefined and heterochromatic states was indicative of increasing noise versus coherent changes in accessibility, we used the ATAC-seq reads to call peaks of localized accessibility. From the WT and hTERT ATAC-seq data from all time points we identified 300,000 ATAC-seq peaks (SData 7) [66]. We did not find a time-dependent decrease in the fraction of reads in peaks during RS. (Figure S12A). We tested experimentally whether the increased heterochromatic accessibility was an artifact of increasing numbers of dead cells by repeating the experiment using late PDL cells with a cross-linking agent (propidium monoazide-PMA) that renders DNA from dead cells inert as previously described [67]. No significant change in read distribution across chromatin states was observed with addition of PMA (Figure S12B).

We then divided the ATAC-seq peaks into each of the 25 discrete chromatin states and calculated the fold change in accessibility for each state and time point compared to the first sample. We observed that changes in peak accessibility mirrors that for all reads. Again we observed a clear increase in accessibility with PDL across the miscellaneous category with a concomitant decrease in all other states in WT WI-38 cells and not in immortalized hTERTs (Figure 5B).

The quiescent or "undefined" state from the miscellaneous category piqued our interest for two reasons, one being that it alone accounted for 20-40% of all ATAC-seq reads and a large 10% increase with senescence. The second point of interest was the observed early onset of change (Figure 5A,B). Dillinger et. al. previously reported that the quiescent state is largely overlapping with both LADs and NADs [62]. Likewise, we found that the quiescent chromatin domains and most of the heterochromatic domains to be markedly gene poor and overwhelmingly overlapping with both LADs and NADs experimentally defined in IMR-90 fibroblasts (Z-score 90 and 150) (Figure S12C). For reference, Figure S12E provides a chromosome level view of the significant overlap between RS accessible ATAC-seq peaks in the quiescent state with NADs and LADs compared to all peaks annotated in WI-38 cells.

We tested and found that ATAC-seq peaks falling within NADs and LADs exhibit increased accessibility with RS ($p < 10$ e-16 for both). Notably, this trend was much greater for peaks in NADs versus peaks in LADs, with a median log2FC of 0.98 for NADs vs. 0.24 for LADs (Figure 5C).

Next, to assess whether overall shifts in NAD/LAD accessibility might be correlated with the changes in RS gene expression, we tested the top up-regulated RS genes and found significant overlap with NADs and LADs. Interestingly, the top RIS or CD induced genes did not exhibit this overlap suggesting this phenomenon is specific to RS (Figure 5D).

Together these results highlight a striking increase in accessibility within nucleolar associated DNA that connects changes in the transcription with a global shift in the epigenome.

2.7 Transcriptional regulators of the RS transcriptome and epigenome

To parse out the regulatory logic of RS gene expression, we leveraged our ATAC-seq data to gain insight into which transcription factors regulate RS accessibility via transcription factor motif analysis. Having cataloged a universe of ATAC-seq peaks with significant changes in accessibility, we next assigned peaks to neighboring genes. Taking the top RS differentially expressed genes, we searched the proximal ATAC-seq peaks enriched transcription factor motifs (Figure 6A). We found enriched motifs for TEAD1, CEBP family transcription factors, SMAD EMT transcription factors, AP1 transcription factors, and multiple FOX family transcription factor motifs.

We also applied an orthogonal gene-independent methodology for determining which motifs are predictive of RS induced increases in ATAC-seq peak accessibility. Consistently, we again found TEAD1 to be the most predictive feature (mean coefficient = 0.25 across 10 models-methods). In addition, we also found evidence for FOXE1 and SMAD1 regulation as well as other senescence related transcription factors, e.g. CEBPB and TP53 (Figure 6B).

Out of the FOX family transcription factors, FOXE1 is unique in that it exhibits one of the most specific and largest increases in RS gene expression (40x). FOXE1 is annotated as a thyroid specific transcription factor with putative roles in thyroid development and cancer [68, 69, 70]. It has also been reported to regulate two signaling molecules up-regulated in the RS time course; TGF-β3 and WNT5A [71, 72]. Furthermore, it has been reported that many of the FOX family transcription factors are relatively promiscuous binders of each others canonical motifs. Thus it is possible that the increased accessibility in peaks with FOX motifs during RS may be driven by FOXE1 activity [73]. Despite the specificity and magnitude of RS induction, FOXE1's function in RS is unclear and warrants further investigation.

We returned to our bulk RNA-seq to test for enrichment of protein-DNA binding events (mapped by ENCODE) in regulatory elements proximal to our RS differentially expressed genes by using Landscape In Silico deletion Analysis (LISA) [74]. Plotting transcription factor enrichment for genes depleted with RS against genes induced during RS revealed 3 broad cohorts of transcription factors; proliferating cell transcription factors, RS transcription factors, and transcription factors whose bind-

bioRxiv preprint doi: https://doi.org/10.1101/2021.05.03.442497; this version posted May 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

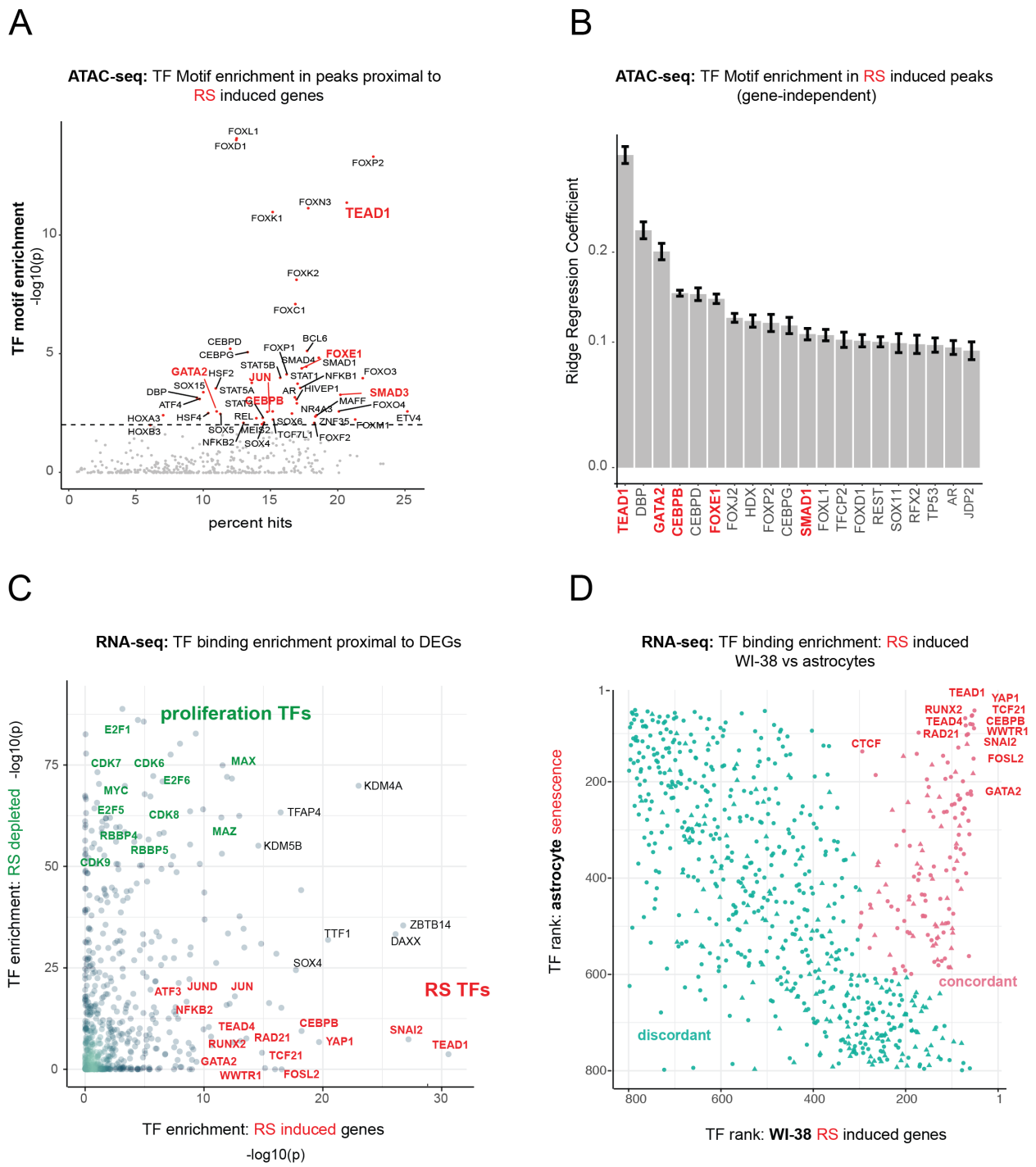

Figure 6: **Master Transcriptional Regulators of Replicative Senescence (A)** Scatter plot of transcription factor motif enrichment in ATAC-seq peaks surrounding the top 1,000 genes from differential expression analysis of RS. y-axis is the -log10 p-value; x-axis is the percent of genes with transcription factor motif. transcription factors of interest are highlighted in red. **(B)** Bar graph of ridge regression coefficient of motif predictive power in model of increasing peak accessibility with RS. transcription factors of interest are highlighted in red. **(C)** Scatterplot of transcription factors enriched for binding in regulatory regions around RS depleted genes (y-axis) vs. RS induced genes(x-axis). Curated cell cycle transcription factors are colored in green; transcription factors of interest e.g. EMT/AP1/YAP1/TEAD1 etc. are colored in red. **(D)** Scatterplot of enriched transcription factors rank for binding enrichment in regulatory regions around senescence induced genes in astrocytes (y-axis) vs. RS induced genes in WI-38 cells (x-axis). transcription factors with discordant ranks/enrichment are colored in turquoise,transcription factors with concordant ranks/enrichment are colored in red; transcription factors of interest are labeled in red.

bioRxiv preprint doi: https://doi.org/10.1101/2021.05.03.442497; this version posted May 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

ing was enriched around both sets of genes (Figure 6C, SData 8). As expected, proliferation specific transcription factors are replete with cell cycle specific transcription factors e.g. E2F and RBP family transcription factors. A large portion of the transcription factors exhibiting RS specificity belong to 4 categories; inflammation transcription factors (NFKB, CEBPB), AP1 sub-units (JUN,JUND,FOSL2), YAP1-TEAD1 components (TEAD1/4, YAP1, WWTR1), and EMT transcription factors (SNAI2, TCF21).

TEAD1 is a member of the TEA domain transcription factors whose functions range across a wide swath of biology depending on context and binding partner. TEAD transcription factors cannot induce gene expression without a cofactor, which is most often YAP1 (yes-associated protein 1) a key downstream effector of Hippo signaling. [75, 76, 77, 78, 79, 80, 81, 82, 83]. Consistent with our identification of TEAD1 and YAP1 as RS regulators, YAP1 activation has been tied to EMT, anti-apoptosis, telomere dysfunction, inflammation, and positive regulation of fatty acid oxidation [84, 85, 86, 87, 88].

Lastly, given that YAP1/TEAD1 activity appears to increase during RS against the backdrop of an altered epigenomic context, we tested for an interaction between TEAD1 motifs and the increasingly accessible NADs and LADs. We found that not only are TEAD1 sites significantly enriched within NADs and LADs (($p < 7.17e-07$, and $p < 6.12e-03$ respectively hypergeometric), but we also discovered a greater-than-additive increase in accessibility with RS for TEAD1 motifs that occur within a NAD ($p < 3.12$ e-07). These results suggest that there is a functional connection between TEAD1 activity and NAD domains during RS. The changing epigenetic context of RS may alter TEAD1 binding and thus target activity with increasing PDL.

Collectively, these analyses uncover a common theme amongst putative regulatory transcription factors; Hippo signaling (YAP1/TEAD1), EMT transcription factors, and TGF-*B* signaling (SNAI2, SMAD activity). SNAI2 has been shown to work in tandem with the YAP1/TEAD1 complex and these pathways often work towards similar biological ends [89, 84]. Together, these transcription factors are reported as highly involved with proliferation, EMT, ECM production, fibrosis, and apoptosis avoidance [90].

Lastly, given that the LISA results are based on binding events collected from a vast multitude of cell lines, we wondered if the same transcription factors might be found to regulate senescence in a completely different cellular context. To test this we took significantly induced genes from a senescence model using astrocytes as the cellelar context and oxidative stress as the senescence trigger. [91]. Plotting transcription factor enrichment for genes induced in astrocyte senescence against transcription factor enrichment for WI-38 RS genes revealed that although there were a substantial number of discordant transcription factors, there was a clear population of transcription factors highly enriched in both senescence models (Figure 6D). Notably, the top concordant transcription factors ranked at the top in both contexts and recapitulate all the previous results e.g. YAP1/TEAD1, SNAI2, CEBP family transcription factors and AP1 subunits.

2.8 scRNA-seq trajectory analysis resolves WI-38 cells' approach to replicative senescence

A wealth of recent work in single cell transcriptomics has demonstrated that ordering single cells in a process-specific trajectory often reveals nuanced timing and dynamics of gene expression that bulk assays cannot capture [92, 93]. Mapping this trajectory is frequently referred to as "pseudotime analysis". We employed pseudotime analysis to arrange single WI-38 cells along a pathway to senescence (SData 3). As expected, early and late PDL cells concentrated at the beginning and end of the pseudotime trajectory respectively (Figure 7A). We next performed differential expression analysis to identify genes that change significantly over pseudotime. We plotted examples of genes changing early (CENPK–an S-phase cell-cycle-regulated gene), midway (SNAI2–a master regulator of EMT), or late (PAPPA–a prominent SASP factor) in pseudotime in Figure 7B.

We next generated gene expression trajectories for the top 5,000 differentially expressed genes across pseudotime (Figure S13, SData 9).To identify the temporal relationship between the biological processes and transcription factors that compose RS in WI-38 cells, we clustered and ordered pseudotime trajectories (SData 9). Broadly, we classified the pseudotime expression pattern as early, transition, or late based on the maximum median value of all constituent genes for each cluster (Figure 7C). We performed LISA transcription factor and GO enrichment analysis on each cluster and used our findings to assign putative functional labels across pseudotime (Figure 7D, Figure S14D, Sdata 10).

Early pseudotime is dominated by the transcription factors (E2F) and GO terms associated with cell cycle progression through the G2M and S phases. Moving down the y-axis deeper into pseudotime, we observe that the next primary cluster of transcription factors and functional annotations exhibited widespread enrichment across all of pseudotime. Furthermore, transcription factors enriched in this cluster are involved with basic cellular functions (e.g. euchromatin maintenance, transcription, growth) and are likely representative of normal WI-38 function in G1 phase.

Moving forward into the transition phase of pseudotime, we observed an enrichment of transcription factors regulating higher order chromatin structure (CTCF) and epigenetic silencing (polycomb group complex). Also in the transition region of pseudotime, the earliest enrichments appear for transcription factors related to EMT (SNAI2), TGF-β signaling (SMAD3), YAP1/TEAD1 activity, and the AP1 complex (JUN, FOS, FOSL2). We observed continued enrichment for these transcription factors and processes throughout the rest of the transition phase and into late pseudotime.

Lastly, and late in pseudotime, we observed enrichment of transcription factors and functional annotations related to regulation of inflammatory processes (NFKB1, RELA, ERG1, CEBPB) and changes in cellular morphology (Figure 7D, Figure S14). These observations are consistent with observations made with bulk RNA-seq that RS WI-38 cells exhibit transcriptomic features similar to that observed in TGF-β signaling, EMT, and

bioRxiv preprint doi: https://doi.org/10.1101/2021.05.03.442497; this version posted May 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

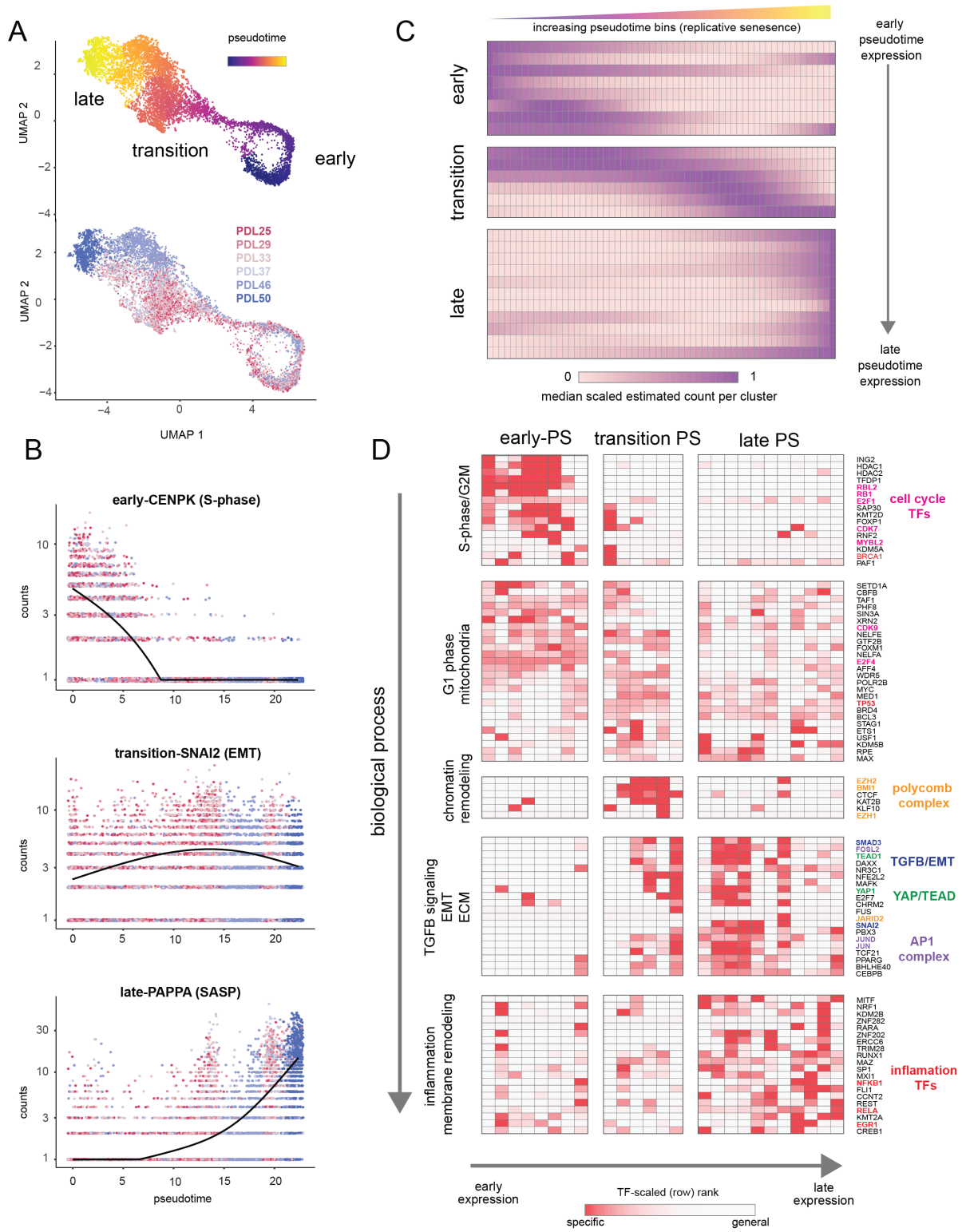

Figure 7: **Pseudotime (PS) analysis of WI-38 approach to replicative senescence using single cell RNA-seq.** **(A)** UMAP projection of single WI-38 cells collected at increasing PDLs (PDL25-red) to (PDL50-blue) and colored by pseudotime (top). **(B)** Scatterplot of single cell gene expression across pseudotime with 3 genes representative of changes that occur early in pseudotime (top), during the transition phase (middle) and in late pseudotime (bottom). Each point is a cell colored by PDL–PDL25 (red) to PDL50 (blue). x-axis is pseudotime, and y-axis is counts for the gene of interest. Black line is a cubic spline. **(C)** Hierarchical clustering and heatmap of smoothened gene expression trajectories over 60 pseudotime bins (x-axis) of 25 K-median gene expression clusters (y-axis). Value plotted is the scaled (min expression to max expression ; 0-1) median expression of all genes in the cluster. Clusters are divided into three (early, transition, late) pseudotime categories. **(D)** LISA transcription factor enrichment analysis using the 25 clusters from A divided into the same three (early, transition, late) pseudotime categories. transcription factors and clusters are further divided on vertical axis into putative groupings based on transcription factor functions and GO term enrichment (sFig9-10).

bioRxiv preprint doi: https://doi.org/10.1101/2021.05.03.442497; this version posted May 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

with YAP1/TEAD1 activity. Collectively, these results present a possible order of operations for RS progression that highlights an initial cessation of active mitotic cycling, followed by an epigenetic shift that precedes a strong EMT/TGFβ signal before segueing into a pro-inflammatory secretory state.

2.9 Replicatively senescent WI-38 fibroblasts express canonical myofibroblasts markers and metabolic features

Given the repeated observations linking RS with EMT, (Figure 1,Figure 3), TGF-β and YAP1/TEAD1 activity (Figure 6,Figure 7), we considered the possibility that these processes are connected through the Fibroblast to Myofibroblast Transition (FMT), a subtype of EMT in which stressed or injured fibroblasts differentiate into myofibroblasts [94, 95, 96]. Upon receiving cues mediated by injury or stress (e.g. activated TGF-β), fibroblasts can trans-differentiate into myofibroblasts, whose functions as "professional repair cells" include increased proliferation, migration, apoptosis avoidance, and ECM/collagen deposition to promote tissue repair and wound closure [97, 98].

Previous work has demonstrated that there exists mechanistic and functional association between telomerase inhibition, senescence, and myofibroblasts. Senescence is an integral part of the wound healing processes; upon injury resolution, activation of a senescence-like phenotype prevents unchecked collagen secretion and fibrosis by preventing myofibroblast proliferation and earmarking them for subsequent immune clearance [8, 99, 100, 101, 102, 103, 104].

To further explore this proposition, we retrieved canonical myofibroblast marker genes and direct transcriptional targets of the YAP1/TEAD1 complex to determine to what extent these genes are expressed in WI-38 RS cells at both the RNA and protein level. First, we examined canonical myofibroblast markers [97], collagens produced by myofibroblasts [105], genes upregulated in myofibroblasts derived from idiopathic pulmonary fibrosis patients [106], and effectors and targets of TGF-β signaling across three data modalities: bulk RNA-seq, single cell RNA-seq, and bulk proteomics (Figure 8A).

For the majority of genes in the curated myofibroblast panel, expression increased with PDL and pseudotime. Importantly, expression of smooth-muscle actin (ACTA2), a classic myofibroblast marker, increases strongly in RS but not RIS, and decreases markedly in CD. The expression of follistatin-like protein (FSTL1), also known to be strongly expressed in smooth muscle, shows a similar pattern, as does fibrillin (FBN1). All three of these are associated with smooth muscle and TGF-β family regulation. In addition we observed an increase in both fibrillar and basal lamina collagens in our data at both the RNA and protein levels.

Collagen processing is a multi-step process requiring the coordination of multiple enzymes and metabolites [107]. Review of the collagen synthesis pathway alongside our metabolomic and proteomic data provide further confirmation that WI-38 RS exhibit altered collagen metabolism (Figure 8B). We observed up-regulation of multiple pathway enzymes as well as increased abundance of hydroxyproline, a primary constituent amino acid of collagen protein. In addition, we observe a striking depletion of ascorbate (vitamin C) which is required for proline hydroxylation and is an essential vitamin. It is possible that our observations underestimate the collagen production potential of RS WI-38 cells as they appear limited in terms of collagen production by the amount of supplemented vitamin C [108].

Mellone et al. recently reported that although senescent fibroblasts share features with myofibroblasts, this resemblance does not extend to fibrogenic ECM components, e.g. collagens [101]. However, it is important to note that Mellone et al. focused on RIS rather than replicative senescence, and we similarly observed less or no induction of many of these same genes in our RIS condition (Figure 8A) revealing another important distinction between RS and RIS.

Moving forward in our myofibroblast panel, we observed that expression of TGF-β cytokine, TGF-β1, decreased significantly with RS in both RNA and protein. However, we see robust induction of the TGF-β isotype 2 (TGF-β2) cytokine with RS (Figure 8A) as TGF-β1 abundance drops, which indicates a switch in TGF-β isotypes with RS. It has been shown TGF-β2 is a more potent inducer of the endothelial to mesenchymal transition (EndMT) *in vitro* compared to TGF-β1 and TGF-β3 in human microvascular endothelial cells, and TGF-β2 may be playing a similar role here in inducing FMT and RS in WI-38 cells [109]. The distinct functional roles of different TGF-β isotypes are largely unknown, although both are known to activate the SMAD transcription factors. On the basis of our data, it seems likely that the TGF-β paralog relevant here might not be TGF-β1, but TGF-β2.

2.10 Expression of YAP1/TEAD1 targets during RS

Next, we retrieved a gene set of YAP1/TEAD1 targets assembled by Kurrpa et al. from five separate studies [84]. Taking the intersection of the five YAP1/TEAD1 gene target lists, we only kept genes present in at least 2 of the studies and plotted the remainder in Figure 8C across the three data types. We arranged the YAP1/TEAD1 targets by the two predominant expression patterns–decrease with RS (top) and increase with RS (bottom). As with myofibroblast markers, the data from all three modalities is largely concordant.

Interestingly, we noted a striking bifurcation whereby YAP1 targets tend towards either strong down-regulation or strong up-regulation. The down regulated partition is heavily enriched for classic cell cycle regulated genes such as TOP2A, CDC20, BIRC5, and CDK9 suggesting that this dichotomy in YAP1 activity is heavily influenced by the cell cycle and consistent with recent work [110, 111].

Excitingly, TGF-β2 was one of the two genes found in four out of five collected YAP1/TEAD1 gene sets, supporting the inference that the TGF-β2 and not TGF-β1 is the relevant paralog and potentially regulated by YAP1/TEAD1 in the RS context. In addition, another YAP1 target, thrombospondin-1 (THBS1), is also upregulated with RS; THBS1 is known to act as an activator of TGF-β signaling, and specifically TGF-β2, through proteolytic cleavage of latent TGF-β [112]. Figure S15

bioRxiv preprint doi: https://doi.org/10.1101/2021.05.03.442497; this version posted May 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

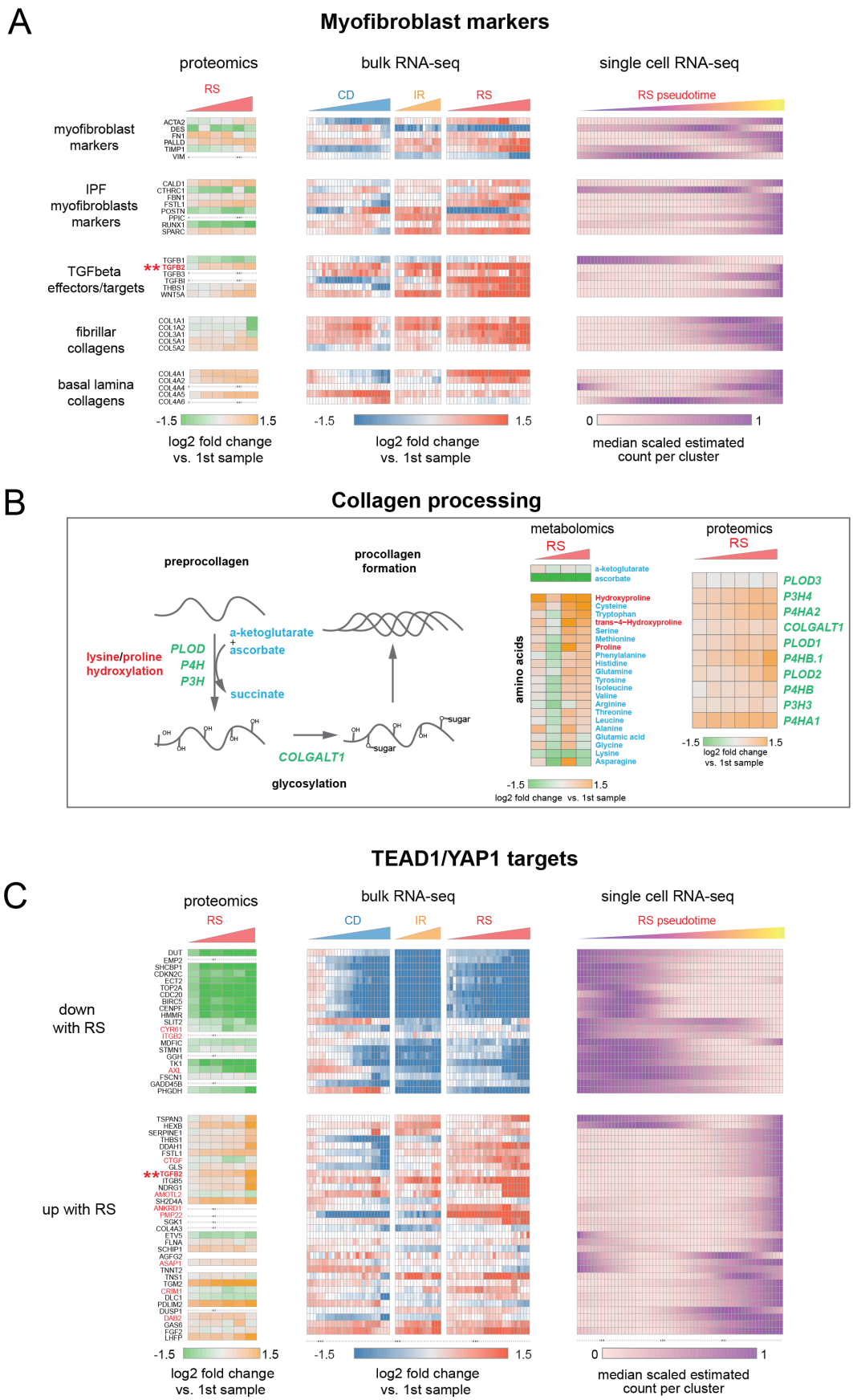

Figure 8: **Expression of Myofibroblast markers and YAP1/TEAD1 targets during RS across multiple data modalities. (A)** Heatmaps of selected genes based on myofibroblast markers and biology from RNA-seq, single cell RNA-seq and proteomics. Values plotted are log2 fold change of each time point versus first. **(B)** Diagram of collagen processing (left) with metabolites in blue and proteins in italicized green. Heatmap of log2 fold changes for metabolites and proteins involved in collagen processing (right). **(C)** Heatmaps of YAP1 targets collected Kurppa et al. across RNA-seq, single cell RNA-seq and proteomics. Values plotted are log2 fold change of each time point versus first. Gene names on the left are colored by the number of studies the gene was present in;>1 black, >2 =red.

bioRxiv preprint doi: https://doi.org/10.1101/2021.05.03.442497; this version posted May 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

displays examples of expression for several of the myofibroblast markers and YAP1 targets in single cell UMAP projections.

In all, the data reveal that RS WI-38 cells share multiple defining transcriptomic and proteomic features with myofibroblasts. Furthermore, we found that a subset of YAP1/TEAD1 targets are both induced with RS, and are principal components of TGF-β signaling (TGF-β2, THBS1). Thus the YAP1/TEAD1 complex may be acting in convergence with TGF-β signaling with increasing PDL to enact a myofibroblast-like state that we recognize as RS.

3 Discussion

The study of RS in human tissue culture has proven to be an informative model for learning how genetic and environmental factors impact cellular senescence. However, the field has not fully taken advantage of the 'omics' revolution. Rekindled by the advent of senolytics, interest in the molecular underpinnings of RS has burgeoned in recent years as researchers seek to design therapeutic strategies for ablating senescent cells [113, 114, 115, 116]. However, many such studies span only one or two systematic data modalities. Here, to fully leverage the power of recent advances in high-dimensional profiling, we revisit the original Hayflick limit in WI-38 lung fibroblasts cells with a battery of assays including RNA-seq, ATAC-seq, scRNA-seq, proteomics, and metabolomics in an effort to capture the defining features of RS at every step of the central dogma and beyond. Our results are summarized graphically in Figure 9.

The kinetics and precise timing of the RS process have been obscured by low temporal resolution and ensemble measurements that cannot differentiate between global shifts in gene expression versus changing proportions of senescent cells [36, 117, 118, 119, 32, 33, 34, 35]. Here, with a combination of high time resolution and single cell RNA-seq, we provide evidence that the early manifestation of the RS gene expression reflects gradual changes on a per cell basis rather than changing cell proportions. In effect, individual cells "show their age" with increasing PDL long before permanently exiting the cell cycle and transiting fully into the senescent state. The implications of this conclusion extend to organismal aging. For example, the percentage of senescent cells calculated from aging organisms varies greatly depending on the marker/phenotype used (reviewed in [120]. The reported disparities could be explained in part by the use of early versus end stage markers. Likewise, it is possible the reported increase in fibroblast heterogeneity and altered functionality with age is a direct result of cells slowly moving along a spectrum towards RS [121, 122]. Importantly, the gradual progression suggests that cells need not reach the endpoint to elicit a phenotype. For instance, proliferative fibroblasts isolated from IPF patients exhibited multiple senescent features and phenotypes in addition to accelerated senescence progression [123]. Lastly, this phenomenon is not constrained to fibroblasts as we observe that the salient regulatory features of RS extend to cell types as distant as astrocytes (Figure 6D).

In our data, the pattern of gene expression annotated to EMT as a unique feature of RS that consistently presents early and robustly at both the RNA, protein, and single cell level. Given the fundamental nature of the EMT transition with respect to cellular function (development, fibrosis, and wound healing), it is not surprising that this hallmark tracks with multiple proteins and gene sets suggesting drastic metabolic rewiring. In our metabolic data we highlight shifts in carbon and fatty acid utilization that have been reported previously as hallmark metabolic features of EMT. These metabolic changes demonstrate that our observations represent an authentic change in cellular state as opposed to a superficial uptick in a few EMT related genes.

The data presented above across multiple data modalities to provide a clear connection between RS cells and myofibroblasts supported by independent observations at the level of DNA, RNA, protein, transcription factor activity and metabolism (Figure 9). In light of these, RS resembles a specialized subtype of EMT specific to the trans-differentiation of fibroblasts into myofibroblasts (FMT) in response to wound healing [124, 98, 97]. We hypothesize that during fibrotic disease states and/or age, fibroblasts migrate to sites of micro-injuries. As these proliferating fibroblasts become replicatively aged, they are triggered (by DNA damage or other insults) to rewire their metabolism to induce FMT via active epigenetic reorganization (NNMT-SAM/NAD sink). It is important to note here that the *in vitro* DNA damage here arises primarily from telomere erosion documented by the observation that our hTERT control cultures do not exhibit the same changes. However, genotoxic stress *in vivo* may originate from a variety of endogenous and environmental sources e.g. reactive oxygen species, replication stress, chemical exposure etc.

Following increases in DNA accessiblity, expression of newly opened TEAD1/YAP1/SMAD target genes cement FMT transition by promoting fibrosis and a myofibroblast-like, ECM-secreting state. This model is supported by the synergy between TEAD1 motifs and NAD domains we report and reconciles conflicting reports that YAP1/TEAD1 inhibition can both prevent and promote senescence [125, 84, 126, 127]. Basically, the functional consequences of YAP1/TEAD1 inhibition will depend on the epigenetic organization of the cells used.

Further metabolic changes (hexoasmine/collagen synthesis and fatty acid oxidation) then support the new pro-fibrotic state. Finally, end point RS cells reinforce the senescent state and contribute to neighboring cell RS progression via secretion of inflammatory factors and SASP.

We think that this transition is distinct from classic FMT as the endpoint cells are not proliferative, but instead bear striking resemblance to lingering senescent myofibroblasts that can persist long after wound repair is complete [97]. Rather than a privileged or unique state, perhaps RS is better categorized as a DNA damage mediated path to a potentially common stress-induced endpoint.

The complex interplay between metabolism and epigenetic regulation preclude easy determination of a causal factor in translating DNA damage into the RS/EMT program in WI-38 cells. Does EMT regulate metabolism or vice versa? Here we present compelling evidence on the side of metabolism. We

bioRxiv preprint doi: https://doi.org/10.1101/2021.05.03.442497; this version posted May 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

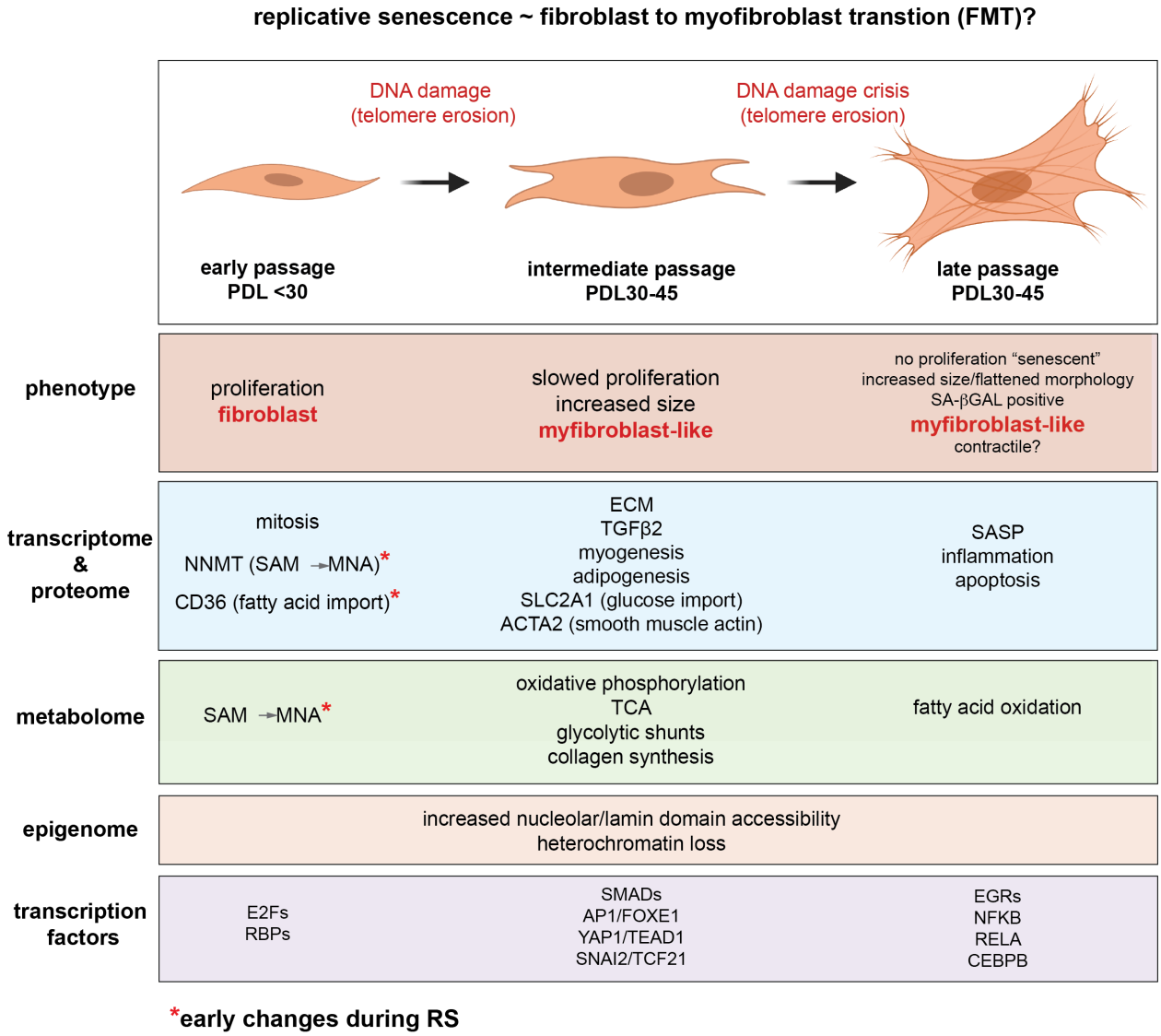

Figure 9: **RS Fibroblast to myofibroblast transition (FMT) model. We divided RS progression into three major categories and summarized our results across all data modalities focusing on features in common with myofibroblasts.**

bioRxiv preprint doi: https://doi.org/10.1101/2021.05.03.442497; this version posted May 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

observe both an early and sharp rise in NNMT expression and activity in addition to a global increase in heterochromatin accessibility. These results are consistent with NNMT's reported role as global epigenetic regulator through its methylation sink activity [58, 59, 128, 57].

Importantly, the observed shifts in repressed chromatin induced by NNMT are functional and may play a central role in fibroblast biology and stress response in multiple contexts. First, Eckert et. al. recently demonstrated that NNMT activity and the resulting heteromchromatin reorganization initiate the expression program of cancer-associated fibroblasts (CAFs) associated with oncogenic stroma *in vivo* [59]. Similar to RS fibroblasts and myofibroblasts, the defining features of CAFs are increased cytokine production, metabolic rewiring, and ECM alteration and production [129]. Second, NNMT is one the most up-regulated genes in a TGF-β mediated *in vitro* FMT conversion in WI-38 cells [130]. Finally, the changes in silenced chromatin we observed in RS overlap with induction of gene expression driving the RS phenotype (Figure 5D). In fact, pseudotime analysis argues that chromatin reorganization may precedes the FMT induction as evidenced by the loss of polycomb activity that appears prior to enrichment of the EMT and YAP1/TEAD1 transcription factors enrichments (Figure 7D).

The path to replicative senescence process has many *in vivo* parallels with implications for aging and pathogenesis. After observing induction in WI-38 cells of IPF myofibroblast markers (Fig. 8A), we expanded our literature search and found a striking overlap (TGF-β signaling, YAP1 activity, and EMT) between our RS data and scRNA-seq expression profiles from alveolar epithelial cells collected from IPF patients [131]. These studies also report a large induction of TGF-β2 relative to TGF-β1 which is consistent with our findings and highlights a clear connection between *in vitro* RS and an *in vivo* disease state.

Another example of such a connection arises from our LISA analysis: one of the top enriched transcription factors, TCF21, has been implicated in atherosclerotic disease progression. Wirka et. al. found that TCF21 promotes the transition of vascular smooth muscle cells into a novel fibroblast-like cell type they dub "fibromyocytes" owing to their possession of both fibroblast and myocyte phenotypes in atherosclerotic lesions in both mice and humans [132].

In general, the FMT hypothesis provides a conceptual framework and integrative model for linking the RS multi-modal phenotypes we observed to multiple human age-related diseases. Given the observation that fibrosis and senescence markers correlate with increasing age in multiple tissues, it is possible that FMT might be a widespread phenomenon underlying many age-related pathologies [55, 133]. Future work harnessing multi-modal single cell technology coupled with relevant *in vivo* models will aid greatly in determining the exact order of events and physiological import.

4 Methods

4.1 Cell Culture

4.1.1 WI-38 maintenance for RS time course-WT and hTERT

WI-38 cells were obtained from the Coriell Institute (AG06814-N) at PDL 15. The cells were grown in Dulbecco's modified Eagle's medium (DMEM, Gibco) with 1 mg/ml of glucose, supplemented with 10% dialyzed fetal bovine serum (FBS) and maintained in an incubator set to 37° C, 5% CO2, and 20% O2. The cells were maintained on 10 cm collagen coated plates (Corning) and split every four to five days until cells slowed down in growth (PDL 37) with approach to senescence and were split every seven days instead. No change was made to hTERT sampling. Cells were tested every month for mycoplasma contamination every month using the MycoAlert Mycoplasma Detection Kit (Lonza). A smaller subset of three PDL timepoints (PDL 45, PDL 55, PDL 56) generated in a secondary time course to ensure "deep" senescence (2-4 months after cell cycle cessation) was profiled as previously reported [134].

4.1.2 WI-38 irradiation time course

WI-38 cells at PDL 20 were plated in DMEM supplemented with 10% FBS with 50,000 cells per well in a 6-well collagen coated plate. Cells were allowed to settle for 2 hours and were subsequently treated with 10 Gy of X-rays. Cells were sampled between 1 day and 9 days for transcriptome profiling. RNA was extracted according to methods below.

4.1.3 WI-38 cell density time course

WI-38 cells at PDL 23 were grown on 10 cm collagen coated plates in DMEM supplemented with 10% FBS. Cells were grown and sampled intermittently between 1 day and 10 days for transcriptome profiling. RNA was extracted according to methods below.

4.1.4 SA-βGAL staining

Cells were stained for senescence associated beta-gal using the Senescence β-Galactosidase Staining Kit (Cell Signaling) by following the manufacturer's published protocols exactly.

4.1.5 WI-38 hTERT cells and lentiviral transduction

293T cells were transfected with the appropriate target plasmid and packaging constructs (lentiviral or retroviral) overnight; 48 hr later, viral supernatant was collected. WI-38 cells were transduced in the presence of 5 μg/mL Polybrene and selected for 7 days with a selection drug.

4.2 Bulk RNA-seq methods

4.2.1 RNA collection and library preparation

Total RNA was extracted from cells using the Direct-zol RNA purification kit (Zymo Research) for all bulk RNA-seq time

bioRxiv preprint doi: https://doi.org/10.1101/2021.05.03.442497; this version posted May 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

course experiments. RNA was quantified using the Fragment Analyzer Standard Sense RNA kit, 15 nt (Agilent Technologies, formerly Advanced Analytical). RNA sequencing libraries were prepared as directed using TruSeq® Stranded mRNA Library Prep Kit and the TruSeq® Stranded Total RNA Library Prep Human/Mouse/Rat Library Prep kit (Illumina), with 1000ng of input material. Samples were amplified for 12 cycles of PCR with TruSeq RNA CD Index Plate (Illumina) and pooled into two separate libraries for each method, before running 3nM libraries across 4 lanes for each set (TruSeq mRNA - 4 lanes and Total RNA - 4 lanes) on the HiSeq 4000 (Illumina).

4.2.2 Read processing and quantification

Reads generated from the Illumina HiSeq 4000 were demultiplexed with bcl2fastq based on the barcode sequence of each sample. Average read depth across samples was 50 milllion paired-end reads. Reads were pseudo-aligned and then quantified using Salmon by deploying the mapping based mode using a Salmon generated index based on Hg38 and optimized for single cell RNA-seq cellRanger v3 to ensure accurate comparison between bulk and single cell RNA-seq. [135].

4.2.3 Differential expression analysis

DESeq2 was used for differential analysis of the RNA-seq data [136]. Wald test was used to estimate fold change and significance using the model: time + batch, where time is a numeric variable representing the fraction of time course complete and batch is a categorical variable used only with RS with the addition of the "deep" senescence time points. DEGs were defined as having p-values < 0.001.

4.2.4 Batch correction and hierarchical clustering

For clustering and visualization we corrected the raw RS count table by batch before converting to TPM with a +1 pseudocount and combining with RIS and CD samples using Combat-seq [137]. Significant genes from each condition were concatenated to generate a universe of significant change genes used for Figure 1. Each sample was then converted to log2 fold change vs. the mean of the initial time point. Prior to hierarchical clustering, the log2 fold changes for each gene were row scaled to emphasize differences across the conditions.

4.2.5 Gene set enrichment analysis

For gene set enrichment analysis, we downloaded the msigDB Hallmark gene sets. For each time course (RS, RIS, and CD) we we ranked all genes by the differential expression log2FC across time generated by DESeq2. Then GSEA was performed on the ranked gene set using the R package "fgsea" which is a new implementation of GSEA in R [30, 138, 31]. By default, GSEA tests for enrichment of each gene set in each condition in both directions. We report the -log10 p-value in Figure 1 and use the normalized enrichment score (NES) to assign directionality of the change.

4.3 Single Cell 3' RNASeq Methods

4.3.1 Cell collection

At each time point, singlet hTERT controls and experimental samples were processed with Chromium Single Cell 3' RNAseq kit V2 (10x Genomics) through cDNA amplification cleanup, where they were frozen at -20°C. Once all time points were converted to cDNA, the frozen cDNAs were thawed and batched for library construction. The following modifications were made to the process: Reverse transcription reactions were brought up to volume with DMEM + 10% FBS instead of water, each emulsion targeted 3000 cells (5200 cells loaded), and cDNA was amplified 12 cycles, with 12 cycles of index PCR. There was no hTert control for time point 3, and one of the replicates for PDL 25 dropped out during library construction. Remaining samples were pooled equimolarly and sequenced on a HiSeq4000 with the standard 26,8,0,98 run configuration.

4.3.2 single cell data processing, normalization, scoring, clustering and DEG analysis

The raw single cell reads were demultiplexed by sample using bcl2fastq. Alignment, cell barcode demultiplexing, transcript quantification and sample merging were carried out with CellRanger 3.0 using the hg38 CR 3.0 gene annoation (also used for bulk RNA-seq). Filtered cell by gene matrices were normalized using Seurat 3.0. and SCTransform [139, 140]. Dimension reductionality was carried out with PCA (n = 50). Clusters were defined using the louvain algorithm and cells were visualized with UMAP projection [141, 142, 143]. For cell cycle scoring and phase determination as well as senescence scoring we applied the Seurat implementation of the scoring function as previously described [37].

DESeq2 was used for differential gene expression analysis to identify significantly changing genes within individual clusters as a function of increasing PDL (PDL cluster) (Fig. 3D) using a an input matrix of gene counts by cells per PDL cluster. Briefly, we summarized gene counts across cells for each clusters and used the top 8,000 genes to expedite analysis. For each PDL cluster, we required that >15 representative cells must exist to be considered in analysis. To visualize differentially expressed genes, we converted single cell counts to CPM (+1 pseudocount) and averaged across PDL cluster and calculated the log2 fold change at each PDL cluster against the earliest PDL for that cluster using a concatenated list of all significantly changing genes.

4.3.3 pseudotime analysis

For pseudotime analysis of the single cell data, we used the R package Monocle 2 and Monocle 3 which implements PCA, Leiden clustering, and UMAP prior to partitioning and trajectory analysis. To focus trajectory analysis on the RS progression, we regressed out the S-Phase and G2M scores for each cell calculated using Seurat (see above) and used only the top genes found to significantly change with RS in bulk RNA-seq (350 up and down) for PCA. [93, 92, 144, 145, 146, 141]. We then employed the Monocle3 "graph test" function to isolate

bioRxiv preprint doi: https://doi.org/10.1101/2021.05.03.442497; this version posted May 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

genes that significantly change as a function of pseudotime. Pseudotime estimation was output from monocle3 using the learn graph function for building a trajectory. Smoothed pseudotime trajectories use for Figure 8C were calculated for significantly changing genes (with pseudotime) by binning cells across pseudotime into 60 bins and using a cubic spline to estimate expression at each bin. For each gene, the smoothed trajectory was set from 0 (minimal) to (1) maximal expression. Genes were organized with K-median clustering (k=25) using cosine similarity. For visualization the median expression value for each pseudotime bin and cluster was calculated. Genes from each cluster were fed into LISA TF analysis (below)

4.4 ATAC-SEQ methods

4.4.1 ATAC-seq Library Preparation and Sequencing

Freshly harvested cells were used for all reactions. Briefly, the cell monolayer was washed with PBS, trypsinized with TrypLE (Gibco), resuspended in media, and cells were pelleted. Cells were counted and 100,000 cells were used in each reaction. Cell lysis, DNA transposition, and library construction was adapted from the Omni-ATAC protocol (Corces et al. 2017). Libraries were amplified for 13 total cycles. Sample purification and size selection was done with SPRI Ampure Beads (Beckman Coulter). Libraries were run on the Pippin HT using 2% agarose cassettes for further size selection (Sage Science). Quality of ATAC-seq libraries were assessed with the Agilent Bioanalyzer 2100 with DNA high sensitivity chips (Agilent Technologies). Libraries were sequenced on the HiSeq4000 with paired end sequencing using 2 x 150 bp reads (Illumina).

4.4.2 ATAC-seq data processing

We first trimmed the raw fastq files with Trimmomatic, then using bowtie2 to align the trimmed reads to hg38. After alignment, we used samtools flags (-f 0x02 and -q2 0) to filter for only properly paired and high quality reads. PCR duplicates are removed using picard MarkDuplicates. Finally for each bam file, we adjusted the reads ends by Tn5 offset (+4 on + strand, -5 on -strand).

For peak calling, we created a condition specific peak atlas by pooling all replicates in a specific condition and applied macs2 for peak calling on the pooled bam file with options (-g hs -p 1e-1 –nomodel –shift -37 –extsize 73). In addition, we performed peak calling on each individual replicates as well. Then we performed irreproducible discovery rate analysis on each condition specific peak atlas for each pair of replicates and filter for peaks that are reproducible in at least two replicates (IDR threshold of 0.05). A single accessibility atlas is created by merging condition-specific peak atlas across all conditions. Peak is assigned to nearest gene if it is within 50kb, otherwise it is annotated as intergenic.

Read count is performed using countOverlaps function from R package GenomicRanges [147]. We performed quantile normalization of the count matrix using normalize.quantiles function of R package preprocessCore. DESeq2 normalization didn't work well for our ATAC-seq data since we observed that median accessibility per sample is still a function of read depth after DESeq2 normalization.

Limma is used for differential accessibility analysis [148]. Fold change and p-value are estimated using moderated t-test statistic based on the model: time + condition + time:condition. We performed separate tests for time point 1 versus each of the other time points.

4.4.3 ATAC-seq chromatin state and NAD/LAD analysis

To quantify ATAC-seq reads in chromatin states we retrieved hg38 ENCODE IMR-90 chromatin state labels from [65]. Next we quantified coverage for each instance of all 25 chromatin states genome wide using countOverlaps function from R package GenomicRanges [147]. We performed quantile normalization of the chromatin state count matrix using normalize.quantiles function of R package preprocessCore (Fig. 6A).

RS ATAC-seq peaks were assigned to chromatin states using the findOverlapsOfPeaks function from the R package GenomicRanges [147]. We used only significantly ($p < 0.001$) changing peaks(with RS) from LIMMA analysis. To simplify the overlap of these two sets of genomic intervals, we took only peaks that fell within ("inside") or encompassed a chromatin state annotation ("inside feature"). Peaks encompassing more than one chromatin state interval were discarded (Fig. 6A,B).

For analysis of NAD and LAD domain overlap we collected IMR-90 NAD labels from [62] and IMR-90 LAD labels from [63]. For calculating overlap Z-score between genomic intervals sets (e.g. NADs) and ATAC-peaks by chromatin state Figure **??**B we used the overlapPermTest function from the R package regioneR [149] using ATAC-seq peaks overlapping chromatin states that significantly increase with RS against a universe of all peaks detected within a given chromatin state type to control for NADbias across different chromatin states.

For testing overlaps between NAD/LAD domains and significantly changing genes, we used the overlapPermTest function from the R package regioneR [149] using the top 1000 genes from each bulk RNA-seq condition (RS,RIS, and CD) against a universe of all genes sampled from the same expression distribution to control NADbias across different levels of expression.

4.5 Metabolomics

4.5.1 Extraction of water-soluble metabolites from lung fibroblast cell culture

WI-38 cells were grown to around 70% confluency in 100 mm cell culture dishes, yielding approximately 2 million cells. Twenty-four hours before metabolite extraction, the medium was aspirated, cells were washed with unconditioned medium and then the medium was replaced. For metabolite extraction, cells were washed once in 37°C warm PBS-buffer (vendor?) immediately followed by the addition of 3.5 mL of freezer-cooled (-20°C) LC-MS grade 80:20 MeOH/H2O (Sigma Aldrich). The plates were then held at -20°C for 2 h, then harvested with a sterile cell scraper while at -20°C and transferred to -20°C cold 5 mL centrifuge tubes (Eppendorf Lo-Bind). After centrifuging the cell-extracts in a 4°C centrifuge for 5 min at 2000 x g, the

bioRxiv preprint doi: https://doi.org/10.1101/2021.05.03.442497; this version posted May 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

supernatants were transferred into new cold centrifuge tubes and dried under nitrogen at 4°C. Dried extracts were stored at -20°C.

4.5.2 LC-MS/MS analysis of cell culture extracts

Dried supernatants were resuspended in 200 µL of water containing 1 µg/mL of deuterated lysine and deuterated phenylalanine and 250 ng/mL of deuterated succinate (Sigma Aldrich) as internal standards. For negative ion mode, the resuspended samples were diluted 1:4 in water, for positive ion mode, they were diluted 1:4 in acetonitrile. Samples were then centrifuged at 18000 x g for 5 minutes, the supernatant was moved to HPLC vials and 5 µL was injected for analysis by LC-MS on Vanquish UPLCs (Thermo Scientific) coupled to Q Exactive Plus mass spectrometers (Thermo Scientific).

For analysis in negative ion mode, separation of compounds was achieved by reverse-phase liquid chromatography on a Zorbax Extend C18 column (150 x 2.1 mm, 1.8 µm particle size) from Agilent. Mobile phase A was 10 mM Tributylamine and 15 mM Acetic Acid in 97:3 water:methanol at pH 4.95 and mobile phase B was methanol. Prior to injection, the column was equilibrated in 0% B for 4 minutes. The gradient eluted isocratically in 0% B for 2.5 minutes, increased to 20% B over 2.5 min, held at 20% B for 2.5 min, increased to 55% B over 5.5 min, increased to 95% B over 2.5 minutes, maintained at 95% B for 3 min, then decreased to 0% B over 0.5 min, where it was held for 3 min, resulting in a total run time of 26 min. The LC separation was carried out at 30°C column temperature using a 200 µL/min flow rate. For MS-analysis, parameters on MS1 were set to 70,000 resolution with an AGC target of 1e6 at a maximum IT of 100 ms. The scan range was 70 to 1050 m/z. MS2 parameters were set to 17,500 resolution at loop count 6, an AGC target of 1e5 at a maximum IT of 50 ms, an isolation window of 1 m/z and an underfill ratio of 1

For analysis in positive ion mode, compounds were separated via hydrophilic liquid interaction chromatography (HILIC), using a SeQuant ZIC-pHILIC column (150 x 2.1 mm, 5 µm particle size) from Millipore. Mobile phase A consisted of 20 mM ammonium bicarbonate at pH 9.2 in H2O, and mobile phase B was acetonitrile. Prior to injection, the column was equilibrated for 6 minutes in 80% B. The gradient then decreased to 20% B over 20 min, then to 15% B over 2 min, returned to 80% B over 0.5 min and held there for 1.5 min for a total run time of 30 min. The column temperature was 35°C with a flow rate of 150 µL/min. For MS-analysis the MS1 parameters were as described for negative ion mode except the AGC target was 3e6. MS2 parameters were the same with following exceptions: dynamic exclusion was set to 25s with an apex trigger from 3 to 10s. Stepped collision energies were set to 20, 40 and 80% NCE.

4.6 Proteomics methods

4.6.1 Materials and sample preparation, extraction and digestion

LC-MS grade organic solvents, water, and tandem mass tag (TMT) isobaric reagents were purchased from Thermo Fisher Scientific (Waltham, MA). Trypsin was ordered from Promega Corporation (Madison, WI) and Lys-C from Wako Chemicals USA (Richmond, VA). Sep-Pak C18 cartridges were from Waters Corporation (Milford, MA). Unless otherwise stated, all other chemicals were purchased from Sigma-Aldrich (St. Louis, MO).

Cell pellets were resuspended in 450 µL of lysis buffer (75mM NaCl, 3% SDS, 50 mM HEPES, pH 8.5) and lysed by passage through a BD PrecisionGlide 21-gauge needle (20X). The lysate was sonicated for 5 minutes and then centrifuged (5 min, 16,000 x g) to remove cellular debris and the supernatant was collected.

Proteins were reduced with 5 mM dithiothreitol (DTT) for 30 minutes at 56°C with shaking. Proteins were then alkylated with 15 mM iodoacetamide (IAM) for 30 minutes at room temperature (RT) in the dark, and excess IAM was quenched with 5 mM DTT for 20 minutes at RT in the dark. Protein purification was accomplished using a methanol-chloroform precipitation. Briefly, 800 µL methanol, 200 µL chloroform and 600 µL water were sequentially added to 200 µL of cell lysate, with 5 second intervals of vortexing between each addition. The samples were centrifuged for 30 minutes (16,000 x g at 4°C) to induce phase separation and both the top and bottom layers were removed. The precipitated protein pellet was washed with 600 µL methanol, vortexed briefly, then centrifuged for 10 minutes (16,000 x g at 4 °C). The methanol layer was removed and protein pellets were dried at RT for 10 minutes. Protein pellets were resuspended in digestion buffer (8 M urea, 50 mM HEPES, pH 8.5). The urea concentration was diluted to 4 M, then proteins were digested with Lys-C overnight (10 ng/µL, 25 °C, 16 h). The urea concentration was further diluted to 1 M and samples were digested with trypsin (5 ng/µL) for 6 hours at 37°C.

Following digestion, peptides were acidified with trifluoroacetic acid (TFA) to a final concentration of 0.5% TFA. Peptides were desalted using Sep-Pak C18 solid-phase extraction (SPE) columns and samples were eluted sequentially, first with 40% acetonitrile (ACN)/0.5% acetic acid and then 80% ACN/0.5% acetic acid. Eluted peptides were dried in a CentriVap Benchtop Vacuum Concentrator (Labconco, Kansas City, MO) running at 30 °C. Peptide concentrations were measured using the Pierce BCA Protein Assay Kit, then 50 µg aliquots of each samples were dried in the CentriVap for further processing.

4.6.2 Tandem mass tag (TMT) labeling

Dried peptides were resuspended in 50 µL 200 mM HEPES/30% anhydrous ACN, then 200 µg of each TMT tag was added to 50 µg peptides. TMT 131c was used as the 'bridge sample' while the other tags (126, 127n, 127c, 128n, 128c, 129n, 129c, 130n, 130c, 131n) were used to label the individual samples. The TMT reaction was incubated for 1 hour at room temperature with gentle shaking, then quenched with 11 µL 5% hydroxylamine/200mM HEPES. All samples were acidified to a final concentration of 0.5% TFA. A small amount (4 µL) of each labeled sample was combined and desalted using StageTips1 to check TMT ratios and labeling efficiency. The TMT-labeled samples were then combined at a 1:1:1:1:1:1:1:1:1:1:1 peptide ratio into 11-plex samples. The

bioRxiv preprint doi: https://doi.org/10.1101/2021.05.03.442497; this version posted May 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

combined samples were desalted using Sep-Pak C18 cartridges and dried under vacuum.

4.6.3 High pH Reversed-Phase (HPRP) fractionation

The pooled TMT-labeled peptides were fractionated using high pH reversed-phase liquid chromatography on an Agilent 1260 Infinity HPLC equipped with a diode array detector set at 215, 220 and 254 nm wavelengths (Agilent Technologies, Santa Clara, CA). Peptides were separated on an Agilent ZORBAX Extend-C18 column (4.6 mm x 250 mm, 5 µm particle size) running at 500 µl/min at 25°C. Peptides were eluted with a gradient with initial starting condition of 100% buffer A (5% ACN, 10 mM ammonium bicarbonate) and 0% buffer B (95% ACN, 10 mM ammonium bicarbonate). Buffer B was increased to 35% over 60 minutes, then ramped up to 100% B in 6 seconds where it was held for 5 minutes. Buffer B was then decreased to 0% over 6 seconds and held for 10 minutes to re-equilibrate the column to original conditions. The samples were fractionated into 96 fractions, then pooled into 12 fractions as previously described2. The fractions were dried under vacuum and re-suspended in 5% ACN/5% formic acid (FA) for LC-MS/MS analysis.

4.6.4 Proteomics Data Acquisition and Analysis

LC-MS/MS Data Acquisition All samples were analyzed by an Orbitrap Fusion Lumos Tribrid mass spectrometer coupled to an EASY-nLC 1200 (Thermo Fisher Scientific). Peptides were separated using a microcapillary column (100 µm x 250 mm long, filled in-house with Maccel C18 AQ resin, 1.8 µm, 120 Å; Sepax Technologies, Newark, DE) operating at 60 °C with a flow rate of 300 nL/min. Peptides were eluted into the mass spectrometer using a 180 min method, with acetonitrile increasing from 6 to 30% over a 165 min linear gradient in 0.125% formic acid. Mass spectrometry data was collected in data-dependent acquisition (DDA) mode. A high resolution MS1 scan (500-1200 m/z range, 60,000 resolution, AGC 5 x 105, 100 ms max. injection time, RF for S-lens 30) was collected in the Orbitrap, and the top 10 precursors were selected for MS2 and MS3 analysis. Ions were isolated using a 0.5 m/z window for MS2 spectra. The MS2 scan was performed in the quadrupole ion trap (CID, AGC 1 x 104, 30% normalized collision energy, 35 ms max. injection time) and the MS3 scan was analyzed in the Orbitrap (HCD, 60,000 resolution, max. AGC 5 x 104, 250 ms max. injection time, 50% normalized collision energy). The max. cycle time was set at 5 s. For TMT reporter ion quantification, up to 6 fragment ions from each MS2 spectra were selected for MS3 analysis using synchronous precursor selection (SPS).

4.6.5 Proteomics data analysis

An in-house software pipeline was used to process all proteomics data3. Raw files were converted to mzXML files and searched against a composite human UniProt database containing forward and reverse sequences using the Sequest algorithm. MS/MS spectra were matched with fully tryptic peptides from this composite dataset using a precursor ion tolerance of 20 ppm and a product ion tolerance of 0.6 Da. TMT modification of peptide N-termini and lysine residues (+229.162932 Da) and carbamidomethylation of cysteine residues (+57.02146 Da) were set as static modifications. Oxidation of methionine residues (+15.99492 Da) was set as a variable modification. Peptide spectral matches were filtered to a 1% false discovery rate (FDR) using linear discriminant analysis (LDA) as previously described3. Non-unique peptides that matched to multiple proteins were assigned to proteins that contained the largest number of matched redundant peptides sequences using the principle of Occam's razor3.

Quantification of TMT reporter ion intensities was performed by extracting the most intense ion within a 0.003 m/z window at the predicted m/z value for each reporter ion. TMT spectra were used for quantification when the sum of the signal-to-noise for all the reporter ions was greater than 200 and the isolation specificity was greater than 0.754.

Peptide level intensities were then used to fit a Bayesian model for protein quantification as previously described5. Posterior means and variances were extracted from the samples and used to generate all figures.

4.7 Regulatory analysis

4.7.1 LISA TF analysis

Identification of putative regulators of the gene expression changes observed in the bulk RNA-seq experiments and the pseudotime analysis was carried out using LISA [74]; a recent algorithm built to levergae the vast amount of protein-DNA interactions catalogued via ENCODE. We used the GUI hosted at http://lisa.cistrome.org/ to enter lists of genes derived from K-medians clustering (above). The online GUI has a max input of 500 genes. Accordingly, for clusters containing >500 genes, each gene was ranked by correlation with the cluster median and the top 500 genes were used. LISA output consists of a ranked file of transcription factors and chromatin modifiers with enrichment p-values associated with specific ENCODE experiments. Given that the calculated p-values were derived from gene sets of different sizes and from different number of supporting experiments, we used the ranks as the input parameter for the LISA analysis shown in Figures 7A and 8D. A universe of top factors across clusters was compiled by concatenating the top 5 TFs from each individual LISA output. The rank matrix was then centered across clusters to identify TFs with the highest rank in specific clusters.

4.7.2 ATAC-seq peak motif enrichment

We built a binary peak-by-motif matrix where each row is a binary vector representing the presence of 405 motifs in CIS-BP that are expressed in WI-38 (cite CIS-BP). In order to characterize transcription factor activity changes during senescence, we implemented both a gene-centric and a gene-indepedent approach.

In the gene-centric approach, we tested enrichment of each of the 405 motifs in peaks that are 1) associated with a gene set of interest (within 50kb), and 2) significantly associated

bioRxiv preprint doi: https://doi.org/10.1101/2021.05.03.442497; this version posted May 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

with senescence (adjusted p-value < 0.05 in TP7 vs TP1) as compared to all the other peaks using binomial test.

Alternatively, we trained a ridge logistic regression model using the binary motif matrix as features to distinguish peaks of significantly increased accessibility during senescence from peaks of significantly reduced accessibility in TP7 vs TP1. A model trained on two thirds of the data distinguishes the two sets of peaks with AUC=0.67 on held-out peaks. We then trained 10 independent models using all the data to evaluate the coefficients to identify features (motifs) that are most predictive of senescence.

5 Data Availability

Our data have been submitted to the NCBI Gene Expression Omnibus under accession number XXX. We have additionally made our data and software tools available at XXX.research.calicolabs.com.

6 Acknowledgements

The authors would like to thank Ari Firestone, Jason Rogers, Nick Bernstein, Jeff Settleman, Rochelle Buffenstein, Emily Stoops, and Antoine Roux for their feedback on the experiments, analysis and the manuscript. In addition the authors would also like to thank Leonard Hayflick for guidance and discussion.

7 Author Contributions

MR, DB, ADL, and RC conceived the project. MC and DGH planned the experiment. MC, RW, AI, JG, LC, TM, and TV performed the experiments. DGH led the data analysis. DGH, HY, IS, and JO performed data analysis. DGH, MC, DB and DRK wrote and edited the manuscript. MR, DB, ADL, RC, AF, CK, BB, FM, DRK, DGH provided support and oversight for the project.

References

[1] Leonard Hayflick and Paul S Moorhead. THE LIMITED IN VITRO LIFETIME OF HUMAN DIPLOID CELL STRAINS, 1964.

[2] Calvin B Harley, A Bruce Futcher, and Carol W Greider. Telomeres shorten during ageing of human fibroblasts, 1990.

[3] Carlos López-Otín, Maria A Blasco, Linda Partridge, Manuel Serrano, and Guido Kroemer. The hallmarks of aging, 2013.

[4] M Meyerson, C M Counter, E N Eaton, L W Ellisen, P Steiner, S D Caddle, L Ziaugra, R L Beijersbergen, M J Davidoff, Q Liu, S Bacchetti, D A Haber, and R A Weinberg. hEST2, the putative human telomerase catalytic subunit gene, is up-regulated in tumor cells and during immortalization. *Cell*, 90(4):785–795, August 1997.

[5] A G Bodnar, M Ouellette, M Frolkis, S E Holt, C P Chiu, G B Morin, C B Harley, J W Shay, S Lichtsteiner, and W E Wright. Extension of life-span by introduction of telomerase into normal human cells. *Science*, 279(5349):349–352, January 1998.

[6] Juan Carlos Acosta, Ana Banito, Torsten Wuestefeld, Athena Georgilis, Peggy Janich, Jennifer P Morton, Dimitris Athineos, Tae-Won Kang, Felix Lasitschka, Mindaugas Andrulis, Gloria Pascual, Kelly J Morris, Sadaf Khan, Hong Jin, Gopuraja Dharmalingam, Ambrosius P Snijders, Thomas Carroll, David Capper, Catrin Pritchard, Gareth J Inman, Thomas Longerich, Owen J Sansom, Salvador Aznar Benitah, Lars Zender, and Jesús Gil. A complex secretory program orchestrated by the inflammasome controls paracrine senescence. *Nat. Cell Biol.*, 15(8):978–990, August 2013.

[7] Jean-Philippe Coppé, Katalin Kauser, Judith Campisi, and Christian M Beauséjour. Secretion of vascular endothelial growth factor by primary human fibroblasts at senescence. *J. Biol. Chem.*, 281(40):29568–29574, October 2006.

[8] Marco Demaria, Naoko Ohtani, Sameh A Youssef, Francis Rodier, Wendy Toussaint, James R Mitchell, Remi-Martin Laberge, Jan Vijg, Harry Van Steeg, Martijn E T Dollé, Jan H J Hoeijmakers, Alain de Bruin, Eiji Hara, and Judith Campisi. An essential role for senescent cells in optimal wound healing through secretion of PDGF-AA. *Dev. Cell*, 31(6):722–733, December 2014.

[9] Jean-Philippe Coppé, Christopher K Patil, Francis Rodier, Yu Sun, Denise P Muñoz, Joshua Goldstein, Peter S Nelson, Pierre-Yves Desprez, and Judith Campisi. Senescence-associated secretory phenotypes reveal cell-nonautonomous functions of oncogenic RAS and the p53 tumor suppressor. *PLoS Biol.*, 6(12):2853–2868, December 2008.

[10] Vassilis Gorgoulis, Peter D Adams, Andrea Alimonti, Dorothy C Bennett, Oliver Bischof, Cleo Bishop, Judith Campisi, Manuel Collado, Konstantinos Evangelou, Gerardo Ferbeyre, Jesús Gil, Eiji Hara, Valery Krizhanovsky, Diana Jurk, Andrea B Maier, Masashi Narita, Laura Niedernhofer, João F Passos, Paul D Robbins, Clemens A Schmitt, John Sedivy, Konstantinos Vougas, Thomas von Zglinicki, Daohong Zhou, Manuel Serrano, and Marco Demaria. Cellular senescence: Defining a path forward. *Cell*, 179(4):813–827, October 2019.

[11] Alejandra Hernandez-Segura, Jamil Nehme, and Marco Demaria. Hallmarks of cellular senescence. *Trends in cell biology*, 28(6):436–453, 2018.

[12] Judith Campisi and Fabrizio D'Adda Di Fagagna. Cellular senescence: when bad things happen to good cells. *Nature reviews Molecular cell biology*, 8(9):729–740, 2007.

bioRxiv preprint doi: https://doi.org/10.1101/2021.05.03.442497; this version posted May 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

[13] Olga A Sedelnikova, Izumi Horikawa, Drazen B Zimonjic, Nicholas C Popescu, William M Bonner, and J Carl Barrett. Senescing human cells and ageing mice accumulate dna lesions with unrepairable double-strand breaks. *Nature cell biology*, 6(2):168–170, 2004.

[14] Jessie C Jeyapalan and John M Sedivy. Cellular senescence and organismal aging. *Mechanisms of ageing and development*, 129(7-8):467–474, 2008.

[15] Utz Herbig, Mark Ferreira, Laura Condel, Dee Carey, and John M Sedivy. Cellular senescence in aging primates. *Science*, 311(5765):1257–1257, 2006.

[16] Bennett G Childs, Matej Durik, Darren J Baker, and Jan M Van Deursen. Cellular senescence in aging and age-related disease: from mechanisms to therapy. *Nature medicine*, 21(12):1424–1435, 2015.

[17] Kalliopi D Tsakiri, Jennifer T Cronkhite, Phillip J Kuan, Chao Xing, Ganesh Raghu, Jonathan C Weissler, Randall L Rosenblatt, Jerry W Shay, and Christine Kim Garcia. Adult-onset pulmonary fibrosis caused by mutations in telomerase. *Proc. Natl. Acad. Sci. U. S. A.*, 104(18):7552–7557, May 2007.

[18] Mary Y Armanios, Julian J L. Chen, Joy D Cogan, Jonathan K Alder, Roxann G Ingersoll, Cheryl Markin, William E Lawson, Mingyi Xie, Irma Vulto, John A Phillips, Peter M Lansdorp, Carol W Greider, and James E Loyd. Telomerase mutations in families with idiopathic pulmonary fibrosis, 2007.

[19] Jennifer T Cronkhite, Chao Xing, Ganesh Raghu, Kelly M Chin, Fernando Torres, Randall L Rosenblatt, and Christine Kim Garcia. Telomere shortening in familial and sporadic pulmonary fibrosis. *Am. J. Respir. Crit. Care Med.*, 178(7):729–737, October 2008.

[20] Jonathan K Alder, Julian J-L Chen, Lisa Lancaster, Sonye Danoff, Shu-chih Su, Joy D Cogan, Irma Vulto, Mingyi Xie, Xiaodong Qi, Rubin M Tuder, et al. Short telomeres are a risk factor for idiopathic pulmonary fibrosis. *Proceedings of the National Academy of Sciences*, 105(35):13051–13056, 2008.

[21] Bridget D Stuart, Jungmin Choi, Samir Zaidi, Chao Xing, Brody Holohan, Rui Chen, Mihwa Choi, Pooja Dharwadkar, Fernando Torres, Carlos E Girod, et al. Exome sequencing links mutations in parn and rtel1 with familial pulmonary fibrosis and telomere shortening. *Nature genetics*, 47(5):512–517, 2015.

[22] Ganesh Raghu, Derek Weycker, John Edelsberg, Williamson Z Bradford, and Gerry Oster. Incidence and prevalence of idiopathic pulmonary fibrosis. *American journal of respiratory and critical care medicine*, 174(7):810–816, 2006.

[23] Anna Duckworth, Michael A Gibbons, Richard J Allen, Howard Almond, Robin N Beaumont, Andrew R Wood, Katie Lunnon, Mark A Lindsay, Louise V Wain, Jess Tyrrell, et al. Telomere length and risk of idiopathic pulmonary fibrosis and chronic obstructive pulmonary disease: a mendelian randomisation study. *The Lancet Respiratory Medicine*, 2020.

[24] Bridget D Stuart, Joyce S Lee, Julia Kozlitina, Imre Noth, Megan S Devine, Craig S Glazer, Fernando Torres, Vaidehi Kaza, Carlos E Girod, Kirk D Jones, et al. Effect of telomere length on survival in patients with idiopathic pulmonary fibrosis: an observational cohort study with independent validation. *The lancet Respiratory medicine*, 2(7):557–565, 2014.

[25] Jinghong Dai, Hourong Cai, Hui Li, Yi Zhuang, Haiyan Min, Yanting Wen, Jie Yang, Qian Gao, Yi Shi, and Long Yi. Association between telomere length and survival in patients with idiopathic pulmonary fibrosis. *Respirology*, 20(6):947–952, 2015.

[26] Robert J Pignolo, João F Passos, Sundeep Khosla, Tamara Tchkonia, and James L Kirkland. Reducing senescent cell burden in aging and disease. *Trends in molecular medicine*, 2020.

[27] M Borghesan, WMH Hoogaars, M Varela-Eirin, N Talma, and M Demaria. A senescence-centric view of aging: implications for longevity and disease. *Trends in Cell Biology*, 2020.

[28] Alejandra Hernandez-Segura, Tristan V de Jong, Simon Melov, Victor Guryev, Judith Campisi, and Marco Demaria. Unmasking transcriptional heterogeneity in senescent cells. *Curr. Biol.*, 27(17):2652–2660.e4, September 2017.

[29] Arthur Liberzon, Chet Birger, Helga Thorvaldsdóttir, Mahmoud Ghandi, Jill P Mesirov, and Pablo Tamayo. The molecular signatures database hallmark gene set collection. *Cell systems*, 1(6):417–425, 2015.

[30] Aravind Subramanian, Pablo Tamayo, Vamsi K Mootha, Sayan Mukherjee, Benjamin L Ebert, Michael A Gillette, Amanda Paulovich, Scott L Pomeroy, Todd R Golub, Eric S Lander, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proceedings of the National Academy of Sciences*, 102(43):15545–15550, 2005.

[31] Vamsi K Mootha, Cecilia M Lindgren, Karl-Fredrik Eriksson, Aravind Subramanian, Smita Sihag, Joseph Lehar, Pere Puigserver, Emma Carlsson, Martin Ridderstråle, Esa Laurila, et al. Pgc-1α-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. *Nature genetics*, 34(3):267–273, 2003.

[32] Huanyin Tang, Anke Geng, Tengjiao Zhang, Chen Wang, Ying Jiang, and Zhiyong Mao. Single senescent cell sequencing reveals heterogeneity in senescent cells induced by telomere erosion. *Protein & cell*, 10(5):370–375, 2019.

[33] Christopher D Wiley, James M Flynn, Christapher Morrissey, Ronald Lebofsky, Joe Shuga, Xiao Dong, Marc A Unger, Jan Vijg, Simon Melov, and Judith Campisi. Analysis of individual cells identifies cell-to-cell variability following induction of cellular senescence. *Aging cell*, 16(5):1043–1050, 2017.

[34] M Shamim Nassrally, Ashley Lau, Katherine Wise, Noah John, Sanjeev Kotecha, Kar Lai Lee, and Robert F

bioRxiv preprint doi: https://doi.org/10.1101/2021.05.03.442497; this version posted May 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Brooks. Cell cycle arrest in replicative senescence is not an immediate consequence of telomere dysfunction. *Mechanisms of ageing and development*, 179:11–22, 2019.

[35] RG Whitney et al. Intraclonal variation in proliferative potential of human diploid fibroblasts: stochastic mechanism for cellular aging. *Science*, 207(4426):82–84, 1980.

[36] João F Passos, Gabriele Saretzki, Shaheda Ahmed, Glyn Nelson, Torsten Richter, Heiko Peters, Ilka Wappler, Matthew J Birket, Graham Harold, Karin Schaeuble, Mark A Birch-Machin, Thomas B L Kirkwood, and Thomas von Zglinicki. Mitochondrial dysfunction accounts for the stochastic heterogeneity in telomere-dependent senescence. *PLoS Biol.*, 5(5):e110, May 2007.

[37] Sonia Nestorowa, Fiona K Hamey, Blanca Pijuan Sala, Evangelia Diamanti, Mairi Shepherd, Elisa Laurenti, Nicola K Wilson, David G Kent, and Berthold Göttgens. A single-cell resolution map of mouse hematopoietic stem and progenitor cell differentiation. *Blood*, 128(8):e20–31, August 2016.

[38] A Macieira-Coelho and B Azzarone. Aging of human fibroblasts is a succession of subtle changes in the cell cycle and has a final short stage with abrupt events. *Experimental Cell Research*, 141(2):325–332, 1982.

[39] Mikolaj Ogrodnik, Hanna Salmonowicz, Diana Jurk, and João F Passos. Expansion and cell-cycle arrest: common denominators of cellular senescence. *Trends in biochemical sciences*, 44(12):996–1008, 2019.

[40] Gabriel E Neurohr, Rachel L Terry, Jette Lengefeld, Megan Bonney, Gregory P Brittingham, Fabien Moretto, Teemu P Miettinen, Laura Pontano Vaites, Luis M Soares, Joao A Paulo, et al. Excessive cell growth causes cytoplasm dilution and contributes to senescence. *Cell*, 176(5):1083–1097, 2019.

[41] P Marlene Absher, RG Absher, and WD Barnes. Genealogies of clones of diploid fibroblasts: Cinemicrophotographic observations of cell division patterns in relation to population age. *Experimental cell research*, 88(1):95–104, 1974.

[42] Werner Zwerschke, Sybille Mazurek, Petra Stöckl, Eveline Hütter, Erich Eigenbrodt, and Pidder Jansen-Dürr. Metabolic analysis of senescent human fibroblasts reveals a role for amp in cellular senescence. *Biochemical journal*, 376(2):403–411, 2003.

[43] Mikolaj Ogrodnik, Satomi Miwa, Tamar Tchkonia, Dina Tiniakos, Caroline L Wilson, Albert Lahat, Christoper P Day, Alastair Burt, Allyson Palmer, Quentin M Anstee, et al. Cellular senescence drives age-dependent hepatic steatosis. *Nature communications*, 8(1):1–12, 2017.

[44] Amy C Flor, Don Wolfgeher, Ding Wu, and Stephen J Kron. A signature of enhanced lipid metabolism, lipid peroxidation and aldehyde stress in therapy-induced senescence. *Cell death discovery*, 3(1):1–12, 2017.

[45] Hermann Unterluggauer, Sybille Mazurek, Barbara Lener, Eveline Hütter, Erich Eigenbrodt, Werner Zwerschke, and Pidder Jansen-Dürr. Premature senescence of human endothelial cells induced by inhibition of glutaminase. *Biogerontology*, 9(4):247–259, 2008.

[46] Yoshikazu Johmura, Takehiro Yamanaka, Satotaka Omori, Teh-Wei Wang, Yuki Sugiura, Masaki Matsumoto, Narumi Suzuki, Soichiro Kumamoto, Kiyoshi Yamaguchi, Seira Hatakeyama, et al. Senolysis by glutaminolysis inhibition ameliorates various age-associated disorders. *Science*, 371(6526):265–270, 2021.

[47] Wan Hua, Peter Ten Dijke, Sarantos Kostidis, Martin Giera, and Marten Hornsveld. Tgfβ-induced metabolic reprogramming during epithelial-to-mesenchymal transition in cancer. *Cellular and Molecular Life Sciences*, 77(11):2103–2123, 2020.

[48] Cyril Corbet, Estelle Bastien, Joao Pedro Santiago de Jesus, Emeline Dierge, Ruben Martherus, Catherine Vander Linden, Bastien Doix, Charline Degavre, Céline Guilbaud, Laurenne Petit, et al. Tgfβ2-induced formation of lipid droplets supports acidosis-driven emt and the metastatic spreading of cancer cells. *Nature communications*, 11(1):1–15, 2020.

[49] Miguel C Lucena, Patricia Carvalho-Cruz, Joana L Donadio, Isadora A Oliveira, Rafaela M de Queiroz, Monica M Marinho-Carvalho, Mauro Sola-Penna, Iron F de Paula, Katia C Gondim, Mark E McComb, et al. Epithelial mesenchymal transition induces aberrant glycosylation through hexosamine biosynthetic pathway activation. *Journal of Biological Chemistry*, 291(25):12917–12929, 2016.

[50] Chao Lu and Craig B Thompson. Metabolic regulation of epigenetics. *Cell metabolism*, 16(1):9–17, 2012.

[51] Xiaoyang Su, Kathryn E Wellen, and Joshua D Rabinowitz. Metabolic control of methylation and acetylation. *Current opinion in chemical biology*, 30:52–60, 2016.

[52] Emma L James, James AE Lane, Ryan D Michalek, Edward D Karoly, and E Kenneth Parkinson. Replicatively senescent human fibroblasts reveal a distinct intracellular metabolic profile with alterations in nad+ and nicotinamide metabolism. *Scientific reports*, 6(1):1–15, 2016.

[53] Carlos López-Otín, Lorenzo Galluzzi, José MP Freije, Frank Madeo, and Guido Kroemer. Metabolic control of longevity. *Cell*, 166(4):802–821, 2016.

[54] Rafał Kozieł, Christoph Ruckenstuhl, Eva Albertini, Michael Neuhaus, Christine Netzberger, Maria Bust, Frank Madeo, Rudolf J Wiesner, and Pidder Jansen-Dürr. Methionine restriction slows down senescence in human diploid fibroblasts. *Aging cell*, 13(6):1038–1048, 2014.

[55] Matthew J Yousefzadeh, Jing Zhao, Christina Bukata, Erin A Wade, Sara J McGowan, Luise A Angelini, Michael P Bank, Aditi U Gurkar, Collin A McGuckian, Mariah F Calubag, Jonathan I Kato, Christin E Burd,

bioRxiv preprint doi: https://doi.org/10.1101/2021.05.03.442497; this version posted May 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Paul D Robbins, and Laura J Niedernhofer. Tissue specificity of senescent cell accumulation during physiologic and accelerated aging of mice. *Aging Cell*, 19(3):e13094, March 2020.

[56] Bérénice A Benayoun, Elizabeth A Pollina, and Anne Brunet. Epigenetic regulation of ageing: linking environmental inputs to genomic stability. *Nature reviews Molecular cell biology*, 16(10):593–610, 2015.

[57] Olesya A Ulanovskaya, Andrea M Zuhl, and Benjamin F Cravatt. Nnmt promotes epigenetic remodeling in cancer by creating a metabolic methylation sink. *Nature chemical biology*, 9(5):300–306, 2013.

[58] Motoaki Komatsu, Takeshi Kanda, Hidenori Urai, Arata Kurokochi, Rina Kitahama, Shuhei Shigaki, Takashi Ono, Hideo Yukioka, Kazuhiro Hasegawa, Hirobumi Tokuyama, et al. Nnmt activation can contribute to the development of fatty liver disease by modulating the nad+ metabolism. *Scientific reports*, 8(1):1–15, 2018.

[59] Mark A Eckert, Fabian Coscia, Agnieszka Chryplewicz, Jae Won Chang, Kyle M Hernandez, Shawn Pan, Samantha M Tienda, Dominik A Nahotko, Gang Li, Ivana Blaženović, et al. Proteomics reveals nnmt as a master metabolic regulator of cancer-associated fibroblasts. *Nature*, 569(7758):723–728, 2019.

[60] Henrik Sperber, Julie Mathieu, Yuliang Wang, Amy Ferreccio, Jennifer Hesson, Zhuojin Xu, Karin A Fischer, Arikketh Devi, Damien Detraux, Haiwei Gu, et al. The metabolome regulates the epigenetic landscape during naive-to-primed human embryonic stem cell transition. *Nature cell biology*, 17(12):1523–1535, 2015.

[61] Parisha P Shah, Greg Donahue, Gabriel L Otte, Brian C Capell, David M Nelson, Kajia Cao, Varun Aggarwala, Hazel A Cruickshanks, Taranjit Singh Rai, Tony McBryan, Brian D Gregory, Peter D Adams, and Shelley L Berger. Lamin B1 depletion in senescent cells triggers large-scale changes in gene expression and the chromatin landscape. *Genes Dev.*, 27(16):1787–1799, August 2013.

[62] Stefan Dillinger, Tobias Straub, and Attila Németh. Nucleolus association of chromosomal domains is largely maintained in cellular senescence despite massive nuclear reorganisation. *PLoS One*, 12(6):e0178821, June 2017.

[63] Mahito Sadaie, Rafik Salama, Thomas Carroll, Kosuke Tomimatsu, Tamir Chandra, Andrew R J Young, Masako Narita, Pedro A Pérez-Mancera, Dorothy C Bennett, Heung Chong, Hiroshi Kimura, and Masashi Narita. Redistribution of the lamin B1 genomic binding profile affects rearrangement of heterochromatic domains and SAHF formation during senescence. *Genes Dev.*, 27(16):1800–1808, August 2013.

[64] Jason D Buenrostro, Beijing Wu, Howard Y Chang, and William J Greenleaf. ATAC-seq: A method for assaying chromatin accessibility Genome-Wide. *Curr. Protoc. Mol. Biol.*, 109(1), January 2015.

[65] Jason Ernst and Manolis Kellis. Large-scale imputation of epigenomic datasets for systematic annotation of diverse human tissues. *Nat. Biotechnol.*, 33(4):364–376, April 2015.

[66] Yong Zhang, Tao Liu, Clifford A Meyer, Jérôme Eeckhoute, David S Johnson, Bradley E Bernstein, Chad Nusbaum, Richard M Myers, Myles Brown, Wei Li, et al. Model-based analysis of chip-seq (macs). *Genome biology*, 9(9):1–9, 2008.

[67] David G Hendrickson, Ilya Soifer, Bernd J Wranik, Griffin Kim, Michael Robles, Patrick A Gibney, and R Scott McIsaac. A new experimental platform facilitates assessment of the transcriptional and chromatin landscapes of aging yeast. *Elife*, 7, October 2018.

[68] M Zannini. TTF-2, a new forkhead protein, shows a temporal expression in the developing thyroid which is consistent with a role in controlling the onset of differentiation, 1997.

[69] Rihab Kallel, Salima Belguith-Maalej, Abdelmounaim Akdi, Mouna Mnif, Ilhem Charfeddine, Pere Galofré, Abdelmonaim Ghorbel, Mohamed Abid, Ricard Marcos, Hammadi Ayadi, Antonia Velázquez, and Hassen Hadj Kacem. Genetic investigation of FOXE1 polyalanine tract in thyroid diseases: New insight on the role of FOXE1 in thyroid carcinoma, 2011.

[70] Zheng Ding, Ronghu Ke, Yong Zhang, Youben Fan, and Jianxia Fan. FOXE1 inhibits cell proliferation, migration and invasion of papillary thyroid cancer by regulating PDGFA, 2019.

[71] Isabella Venza, Maria Visalli, Luca Parrillo, Mario De Felice, Diana Teti, and Mario Venza. MSX1 and TGF-beta3 are novel target genes functionally regulated by FOXE1. *Hum. Mol. Genet.*, 20(5):1016–1025, March 2011.

[72] Joana S Pereira, Joana Gomes da Silva, Rute Alexandra Tomaz, António Evaristo Pinto, Maria João Bugalho, Valeriano Leite, and Branca Maria Cavaco. Identification of a novel germline FOXE1 variant in patients with familial non-medullary thyroid carcinoma (FNMTC). *Endocrine*, 49(1):204–214, May 2015.

[73] Julia M Rogers, Colin T Waters, Tom C M Seegar, Sanchez M Jarrett, Amelia N Hallworth, Stephen C Blacklow, and Martha L Bulyk. Bispecific forkhead transcription factor FoxN3 recognizes two distinct motifs with different DNA shapes. *Mol. Cell*, 74(2):245–253.e6, April 2019.

[74] Qian Qin, Jingyu Fan, Rongbin Zheng, Changxin Wan, Shenglin Mei, Qiu Wu, Hanfei Sun, Myles Brown, Jing Zhang, Clifford A Meyer, and X Shirley Liu. Lisa: inferring transcriptional regulators through integrative modeling of public chromatin accessibility and ChIP-seq data. *Genome Biol.*, 21(1):32, February 2020.

[75] Anthony Azakie, Lauren Lamont, Jeffrey R Fineman, and Youping He. Divergent transcriptional enhancer factor-1 regulates the cardiac troponin T promoter. *Am.*

bioRxiv preprint doi: https://doi.org/10.1101/2021.05.03.442497; this version posted May 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

J. Physiol. Cell Physiol., 289(6):C1522–34, December 2005.

[76] Z Chen, G A Friedrich, and P Soriano. Transcriptional enhancer factor 1 disruption by a retroviral gene trap leads to heart defects and embryonic lethality in mice, 1994.

[77] A Benhaddou, C Keime, T Ye, A Morlon, I Michel, B Jost, G Mengus, and I Davidson. Transcription factor TEAD4 regulates expression of myogenin and the unfolded protein response genes during C2C12 cell differentiation, 2012.

[78] Fa-Xing Yu, Bin Zhao, and Kun-Liang Guan. Hippo pathway in organ size control, tissue homeostasis, and cancer, 2015.

[79] André Landin Malt, Julie Cagliero, Kevin Legent, Joël Silber, Alain Zider, and Domenico Flagiello. Alteration of TEAD1 expression levels confers apoptotic resistance through the transcriptional up-regulation of livin. *PLoS One*, 7(9):e45498, September 2012.

[80] A Vassilev, K J Kaneko, H Shu, Y Zhao, and M L DePamphilis. TEAD/TEF transcription factors utilize the activation domain of YAP65, a Src/Yes-associated protein localized in the cytoplasm. *Genes Dev.*, 15(10):1229–1241, May 2001.

[81] B Zhao, X Ye, J Yu, L Li, W Li, S Li, J Yu, J D Lin, C-Y Wang, A M Chinnaiyan, Z-C Lai, and K-L Guan. TEAD mediates YAP-dependent gene induction and growth control, 2008.

[82] Shenghong Ma, Zhipeng Meng, Rui Chen, and Kun-Liang Guan. The hippo pathway: biology and pathophysiology. *Annual review of biochemistry*, 88:577–604, 2019.

[83] Stefano Piccolo, Sirio Dupont, and Michelangelo Cordenonsi. The biology of yap/taz: hippo signaling and beyond. *Physiological reviews*, 2014.

[84] Kari J Kurppa, Yao Liu, Ciric To, Tinghu Zhang, Mengyang Fan, Amir Vajdi, Erik H Knelson, Yingtian Xie, Klothilda Lim, Paloma Cejas, Andrew Portell, Patrick H Lizotte, Scott B Ficarro, Shuai Li, Ting Chen, Heidi M Haikala, Haiyun Wang, Magda Bahcall, Yang Gao, Sophia Shalhout, Steffen Boettcher, Bo Hee Shin, Tran Thai, Margaret K Wilkens, Michelle L Tillgren, Mierzhati Mushajiang, Man Xu, Jihyun Choi, Arrien A Bertram, Benjamin L Ebert, Rameen Beroukhim, Pratiti Bandopadhayay, Mark M Awad, Prafulla C Gokhale, Paul T Kirschmeier, Jarrod A Marto, Fernando D Camargo, Rizwan Haq, Cloud P Paweletz, Kwok-Kin Wong, David A Barbie, Henry W Long, Nathanael S Gray, and Pasi A Jänne. Treatment-Induced tumor dormancy through YAP-Mediated transcriptional reprogramming of the apoptotic pathway. *Cancer Cell*, 37(1):104–122.e12, January 2020.

[85] Choong-kun Lee, Seung-hwan Jeong, Cholsoon Jang, Hosung Bae, Yoo Hyung Kim, Intae Park, Sang Kyum Kim, and Gou Young Koh. Tumor metastasis to lymph nodes requires yap-dependent metabolic adaptation. *Science*, 363(6427):644–649, 2019.

[86] Changmin Peng, Yue Zhu, Wanjun Zhang, Qinchao Liao, Yali Chen, Xinyuan Zhao, Qiang Guo, Pan Shen, Bei Zhen, Xiaohong Qian, et al. Regulation of the hippo-yap pathway by glucose sensor o-glcnacylation. *Molecular cell*, 68(3):591–604, 2017.

[87] Xiao Zhang, Yongxia Qiao, Qi Wu, Yan Chen, Shaowu Zou, Xiangfan Liu, Guoqing Zhu, Yinghui Zhao, Yuxin Chen, Yongchun Yu, et al. The essential role of yap o-glcnacylation in high-glucose-stimulated liver tumorigenesis. *Nature communications*, 8(1):1–15, 2017.

[88] Deepavali Chakravarti, Baoli Hu, Xizeng Mao, Asif Rashid, Jiexi Li, Jun Li, Wen-ting Liao, Elizabeth M Whitley, Prasenjit Dey, Pingping Hou, et al. Telomere dysfunction activates yap1 to drive tissue inflammation. *Nature communications*, 11(1):1–13, 2020.

[89] Yi Tang, Tamar Feinberg, Evan T Keller, Xiao-Yan Li, and Stephen J Weiss. Snail/Slug binding interactions with YAP/TAZ control skeletal stem cell self-renewal and differentiation. *Nat. Cell Biol.*, 18(9):917–929, September 2016.

[90] Cho-Long Kim, Sue-Hee Choi, and Jung-Soon Mo. Role of the hippo pathway in fibrosis and cancer. *Cells*, 8(5), May 2019.

[91] Elizabeth P Crowe, Ferit Tuzer, Brian D Gregory, Greg Donahue, Sager J Gosai, Justin Cohen, Yuk Y Leung, Emre Yetkin, Raffaella Nativio, Li-San Wang, et al. Changes in the transcriptome of human astrocytes accompanying oxidative stress-induced senescence. *Frontiers in aging neuroscience*, 8:208, 2016.

[92] Xiaojie Qiu, Qi Mao, Ying Tang, Li Wang, Raghav Chawla, Hannah A Pliner, and Cole Trapnell. Reversed graph embedding resolves complex single-cell trajectories. *Nat. Methods*, 14(10):979–982, October 2017.

[93] Cole Trapnell, Davide Cacchiarelli, Jonna Grimsby, Prapti Pokharel, Shuqiang Li, Michael Morse, Niall J Lennon, Kenneth J Livak, Tarjei S Mikkelsen, and John L Rinn. The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. *Nat. Biotechnol.*, 32(4):381–386, April 2014.

[94] Sem H Phan. Biology of fibroblasts and myofibroblasts. *Proceedings of the American Thoracic Society*, 5(3):334–337, 2008.

[95] Bram Piersma, Saskia de Rond, Paul M N Werker, Stellar Boo, Boris Hinz, Marike M van Beuge, and Ruud A Bank. YAP1 is a driver of myofibroblast differentiation in normal and diseased fibroblasts. *Am. J. Pathol.*, 185(12):3326–3337, December 2015.

[96] Boris Hinz. Formation and function of the myofibroblast during tissue repair. *Journal of Investigative Dermatology*, 127(3):526–537, 2007.

[97] Boris Hinz and David Lagares. Evasion of apoptosis by myofibroblasts: a hallmark of fibrotic diseases. *Nat. Rev. Rheumatol.*, 16(1):11–31, January 2020.

bioRxiv preprint doi: https://doi.org/10.1101/2021.05.03.442497; this version posted May 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

[98] Andrew A Gibb, Michael P Lazaropoulos, and John W Elrod. Myofibroblasts and fibrosis: Mitochondrial and metabolic control of cellular differentiation. *Circulation Research*, 127(3):427–447, 2020.

[99] Valery Krizhanovsky, Monica Yon, Ross A Dickins, Stephen Hearn, Janelle Simon, Cornelius Miething, Herman Yee, Lars Zender, and Scott W Lowe. Senescence of activated stellate cells limits liver fibrosis. *Cell*, 134(4):657–667, August 2008.

[100] Joon-Il Jun and Lester F Lau. The matricellular protein CCN1 induces fibroblast senescence and restricts fibrosis in cutaneous wound healing. *Nat. Cell Biol.*, 12(7):676–685, July 2010.

[101] Massimiliano Mellone, Christopher J Hanley, Steve Thirdborough, Toby Mellows, Edwin Garcia, Jeongmin Woo, Joanne Tod, Steve Frampton, Veronika Jenei, Karwan A Moutasim, Tasnuva D Kabir, Peter A Brennan, Giulia Venturi, Kirsty Ford, Nicolas Herranz, Kue Peng Lim, James Clarke, Daniel W Lambert, Stephen S Prime, Timothy J Underwood, Pandurangan Vijayanand, Kevin W Eliceiri, Christopher Woelk, Emma V King, Jesus Gil, Christian H Ottensmeier, and Gareth J Thomas. Induction of fibroblast senescence generates a non-fibrogenic myofibroblast phenotype that differentially impacts on cancer prognosis. *Aging*, 9(1):114–132, December 2016.

[102] Neetu Razdan, Themistoklis Vasilopoulos, and Utz Herbig. Telomere dysfunction promotes transdifferentiation of human fibroblasts into myofibroblasts. *Aging Cell*, 17(6):e12838, December 2018.

[103] Tianju Liu, Yasuhiro Nozaki, and Sem H Phan. Regulation of telomerase activity in rat lung fibroblasts. *American journal of respiratory cell and molecular biology*, 26(5):534–540, 2002.

[104] Tianju Liu, Biao Hu, Myoung Ja Chung, Matt Ullenbruch, Hong Jin, and Sem H Phan. Telomerase regulation of myofibroblast differentiation. *American journal of respiratory cell and molecular biology*, 34(5):625–633, 2006.

[105] K Zhang, M D Rekhter, D Gordon, and S H Phan. Myofibroblasts and their role in lung collagen gene expression during pulmonary fibrosis. a combined immunohistochemical and in situ hybridization study. *Am. J. Pathol.*, 145(1):114–125, July 1994.

[106] Maximilian Strunz, Lukas M Simon, Meshal Ansari, Jaymin J Kathiriya, Ilias Angelidis, Christoph H Mayr, George Tsidiridis, Marius Lange, Laura F Mattner, Min Yee, Paulina Ogar, Arunima Sengupta, Igor Kukhtevich, Robert Schneider, Zhongming Zhao, Carola Voss, Tobias Stoeger, Jens H L Neumann, Anne Hilgendorff, Jürgen Behr, Michael O'Reilly, Mareike Lehmann, Gerald Burgstaller, Melanie Königshoff, Harold A Chapman, Fabian J Theis, and Herbert B Schiller. Alveolar regeneration through a krt8+ transitional stem cell state that persists in human lung fibrosis. *Nat. Commun.*, 11(1):3559, July 2020.

[107] Harvey Lodish and S Lawrence Zipursky. Molecular cell biology. *Biochem Mol Biol Educ*, 29:126–133, 2001.

[108] N Boyera, I Galey, and BA Bernard. Effect of vitamin c and its derivatives on collagen synthesis and cross-linking by normal human fibroblasts. *International journal of cosmetic science*, 20(3):151–158, 1998.

[109] Harika Sabbineni, Arti Verma, and Payaningal R Somanath. Isoform-specific effects of transforming growth factor β on endothelial-to-mesenchymal transition. *Journal of cellular physiology*, 233(11):8418–8428, 2018.

[110] Wantae Kim, Yong Suk Cho, Xiaohui Wang, Ogyi Park, Xueyan Ma, Hanjun Kim, Wenjian Gan, Eek-hoon Jho, Boksik Cha, Yun-ji Jeung, et al. Hippo signaling is intrinsically regulated during cell cycle progression by apc/ccdh1. *Proceedings of the National Academy of Sciences*, 116(19):9423–9432, 2019.

[111] Nídia de Sousa, Gustavo Rodríguez-Esteban, Jose Ignacio Rojo-Laguna, Emili Saló, and Teresa Adell. Hippo signaling controls cell cycle and restricts cell plasticity in planarians. *PLoS biology*, 16(1):e2002399, 2018.

[112] Solange M F Ribeiro, Maria Poczatek, Stacey Schultz-Cherry, Matteo Villain, and Joanne E Murphy-Ullrich. The activation sequence of thrombospondin-1 interacts with the latency-associated peptide to regulate activation of latent transforming growth Factor-β, 1999.

[113] Corina Amor, Judith Feucht, Josef Leibold, Yu-Jui Ho, Changyu Zhu, Direna Alonso-Curbelo, Jorge Mansilla-Soto, Jacob A Boyer, Xiang Li, Theodoros Giavridis, et al. Senolytic car t cells reverse senescence-associated pathologies. *Nature*, 583(7814):127–132, 2020.

[114] Monica RP Elmore, Lindsay A Hohsfield, Enikö A Kramár, Lilach Soreq, Rafael J Lee, Stephanie T Pham, Allison R Najafi, Elizabeth E Spangenberg, Marcelo A Wood, Brian L West, et al. Replacement of microglia in the aged brain reverses cognitive, synaptic, and neuronal deficits in mice. *Aging cell*, 17(6):e12832, 2018.

[115] Verena Wagner and Jesús Gil. T cells engineered to target senescence, 2020.

[116] Haig Aghajanian, Toru Kimura, Joel G Rurik, Aidan S Hancock, Michael S Leibowitz, Li Li, John Scholler, James Monslow, Albert Lo, Wei Han, et al. Targeting cardiac fibrosis with engineered t cells. *Nature*, 573(7774):430–433, 2019.

[117] Zhou Xu, Khanh Dao Duc, David Holcman, and Maria Teresa Teixeira. The length of the shortest telomere as the major determinant of the onset of replicative senescence. *Genetics*, 194(4):847–857, August 2013.

[118] Stella Victorelli and João F Passos. Telomeres and cell senescence - size matters not. *EBioMedicine*, 21:14–20, July 2017.

[119] J R Smith and R G Whitney. Intraclonal variation in proliferative potential of human diploid fibroblasts: stochastic mechanism for cellular aging. *Science*, 207(4426):82–84, January 1980.

bioRxiv preprint doi: https://doi.org/10.1101/2021.05.03.442497; this version posted May 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

[120] Cellular aging beyond cellular senescence: Markers of senescence prior to cell cycle arrest in vitro and in vivo.

[121] Wisoo Shin, Nicole L. Rosin, Holly Sparks, Sarthak Sinha, Waleed Rahmani, Nilesh Sharma, Matt Workentine, Sepideh Abbasi, Elodie Labit, Jo Anne Stratton, and Jeff Biernaskie. Dysfunction of hair follicle mesenchymal progenitors contributes to age-associated hair loss. *Developmental Cell*, 53(2):185–198.e7, 2020.

[122] Salah Mahmoudi, Elena Mancini, Lucy Xu, Alessandra Moore, Fereshteh Jahanbani, Katja Hebestreit, Rajini Srinivasan, Xiyan Li, Keerthana Devarajan, Laurie Prélot, et al. Heterogeneity in old fibroblasts is linked to variability in reprogramming and wound healing. *Nature*, 574(7779):553–558, 2019.

[123] Hagai Yanai, Albert Shteinberg, Ziv Porat, Arie Budovsky, Alex Braiman, Rolf Zeische, and Vadim E Fraifeld. Cellular senescence-like features of lung fibroblasts derived from idiopathic pulmonary fibrosis patients. *Aging (Albany NY)*, 7(9):664, 2015.

[124] Alyssa A Lombardi, Andrew A Gibb, Ehtesham Arif, Devin W Kolmetzky, Dhanendra Tomar, Timothy S Luongo, Pooja Jadiya, Emma K Murray, Pawel K Lorkiewicz, György Hajnóczky, et al. Mitochondrial calcium exchange links metabolism with the epigenome to control cellular differentiation. *Nature communications*, 10(1):1–17, 2019.

[125] Lina Fu, Yuqiong Hu, Moshi Song, Zunpeng Liu, Weiqi Zhang, Fa-Xing Yu, Jun Wu, Si Wang, Juan Carlos Izpisua Belmonte, Piu Chan, Jing Qu, Fuchou Tang, and Guang-Hui Liu. Up-regulation of FOXD1 by YAP alleviates senescence and osteoarthritis. *PLoS Biol.*, 17(4):e3000201, April 2019.

[126] Qi Xie, Jing Chen, Han Feng, Shengyi Peng, Ursula Adams, Yujie Bai, Li Huang, Ji Li, Junjian Huang, Songshu Meng, and Zengqiang Yuan. YAP/TEAD–Mediated transcription controls cellular senescence, 2013.

[127] Linglu Jia, Weiting Gu, Yunpeng Zhang, Baoqi Jiang, Xu Qiao, and Yong Wen. Activated Yes-Associated protein accelerates cell cycle, inhibits apoptosis, and delays senescence in human periodontal ligament stem cells. *Int. J. Med. Sci.*, 15(11):1241–1250, July 2018.

[128] Pavlos Pissios. Nicotinamide n-methyltransferase: more than a vitamin b3 clearance enzyme. *Trends in Endocrinology & Metabolism*, 28(5):340–353, 2017.

[129] Erik Sahai, Igor Astsaturov, Edna Cukierman, David G DeNardo, Mikala Egeblad, Ronald M Evans, Douglas Fearon, Florian R Greten, Sunil R Hingorani, Tony Hunter, et al. A framework for advancing our understanding of cancer-associated fibroblasts. *Nature Reviews Cancer*, 20(3):174–186, 2020.

[130] Erin Joanne Walker, Deborah Heydet, Timothy Veldre, and Reena Ghildyal. Transcriptomic changes during tgf-β-mediated differentiation of airway fibroblasts to myofibroblasts. *Scientific reports*, 9(1):1–14, 2019.

[131] Yan Xu, Takako Mizuno, Anusha Sridharan, Yina Du, Minzhe Guo, Jie Tang, Kathryn A Wikenheiser-Brokamp, Anne-Karina T Perl, Vincent A Funari, Jason J Gokey, Barry R Stripp, and Jeffrey A Whitsett. Single-cell RNA sequencing identifies diverse roles of epithelial cells in idiopathic pulmonary fibrosis. *JCI Insight*, 1(20):e90558, December 2016.

[132] Robert C Wirka, Dhananjay Wagh, David T Paik, Milos Pjanic, Trieu Nguyen, Clint L Miller, Ramen Kundu, Manabu Nagao, John Coller, Tiffany K Koyano, et al. Atheroprotective roles of smooth muscle cell phenotypic modulation and the tcf21 disease gene as revealed by single-cell analysis. *Nature medicine*, 25(8):1280–1289, 2019.

[133] M Laura Idda, Waverly G McClusky, Valeria Lodde, Rachel Munk, Kotb Abdelmohsen, Martina Rossi, and Myriam Gorospe. Survey of senescent cell markers with age in human tissues. *Aging*, 12(5):4052–4066, March 2020.

[134] Marco De Cecco, Takahiro Ito, Anna P Petrashen, Amy E Elias, Nicholas J Skvir, Steven W Criscione, Alberto Caligiana, Greta Brocculi, Emily M Adney, Jef D Boeke, et al. L1 drives ifn in senescent cells and promotes age-associated inflammation. *Nature*, 566(7742):73–78, 2019.

[135] Rob Patro, Geet Duggal, Michael I Love, Rafael A Irizarry, and Carl Kingsford. Salmon provides fast and bias-aware quantification of transcript expression. *Nat. Methods*, 14(4):417–419, April 2017.

[136] Michael I Love, Wolfgang Huber, and Simon Anders. Moderated estimation of fold change and dispersion for rna-seq data with deseq2. *Genome biology*, 15(12):550, 2014.

[137] Yuqing Zhang, Giovanni Parmigiani, and W Evan Johnson. ComBat-seq: batch effect adjustment for RNA-seq count data. *NAR Genomics and Bioinformatics*, 2(3), 09 2020. lqaa078.

[138] Gennady Korotkevich, Vladimir Sukhov, and Alexey Sergushichev. Fast gene set enrichment analysis. *BioRxiv*, page 060012, 2019.

[139] Christoph Hafemeister and Rahul Satija. Normalization and variance stabilization of single-cell rna-seq data using regularized negative binomial regression. *Genome Biology*, 20:296, 2019.

[140] Tim Stuart, Andrew Butler, Paul Hoffman, Christoph Hafemeister, Efthymia Papalexi, William M Mauck III, Yuhan Hao, Marlon Stoeckius, Peter Smibert, and Rahul Satija. Comprehensive integration of single-cell data. *Cell*, 177:1888–1902, 2019.

[141] Leland McInnes, John Healy, and James Melville. Umap: Uniform manifold approximation and projection for dimension reduction. *arXiv preprint arXiv:1802.03426*, 2018.

[142] Etienne Becht, Leland McInnes, John Healy, Charles-Antoine Dutertre, Immanuel WH Kwok, Lai Guan Ng,

bioRxiv preprint doi: https://doi.org/10.1101/2021.05.03.442497; this version posted May 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Florent Ginhoux, and Evan W Newell. Dimensionality reduction for visualizing single-cell data using umap. *Nature biotechnology*, 37(1):38–44, 2019.

[143] Vincent D Blondel, Jean-Loup Guillaume, Renaud Lambiotte, and Etienne Lefebvre. Fast unfolding of communities in large networks. *Journal of statistical mechanics: theory and experiment*, 2008(10):P10008, 2008.

[144] Junyue Cao, Malte Spielmann, Xiaojie Qiu, Xingfan Huang, Daniel M Ibrahim, Andrew J Hill, Fan Zhang, Stefan Mundlos, Lena Christiansen, Frank J Steemers, et al. The single-cell transcriptional landscape of mammalian organogenesis. *Nature*, 566(7745):496–502, 2019.

[145] Vincent A Traag, Ludo Waltman, and Nees Jan van Eck. From louvain to leiden: guaranteeing well-connected communities. *Scientific reports*, 9(1):1–12, 2019.

[146] Jacob H Levine, Erin F Simonds, Sean C Bendall, Kara L Davis, D Amir El-ad, Michelle D Tadmor, Oren Litvin, Harris G Fienberg, Astraea Jager, Eli R Zunder, et al. Data-driven phenotypic dissection of aml reveals progenitor-like cells that correlate with prognosis. *Cell*, 162(1):184–197, 2015.

[147] Michael Lawrence, Wolfgang Huber, Hervé Pages, Patrick Aboyoun, Marc Carlson, Robert Gentleman, Martin T Morgan, and Vincent J Carey. Software for computing and annotating genomic ranges. *PLoS Comput Biol*, 9(8):e1003118, 2013.

[148] Matthew E Ritchie, Belinda Phipson, DI Wu, Yifang Hu, Charity W Law, Wei Shi, and Gordon K Smyth. limma powers differential expression analyses for rna-sequencing and microarray studies. *Nucleic acids research*, 43(7):e47–e47, 2015.

[149] Bernat Gel, Anna Díez-Villanueva, Eduard Serra, Marcus Buschbeck, Miguel A Peinado, and Roberto Malinverni. regioner: an r/bioconductor package for the association analysis of genomic regions based on permutation tests. *Bioinformatics*, 32(2):289–291, 2016.

bioRxiv preprint doi: https://doi.org/10.1101/2021.05.03.442497; this version posted May 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

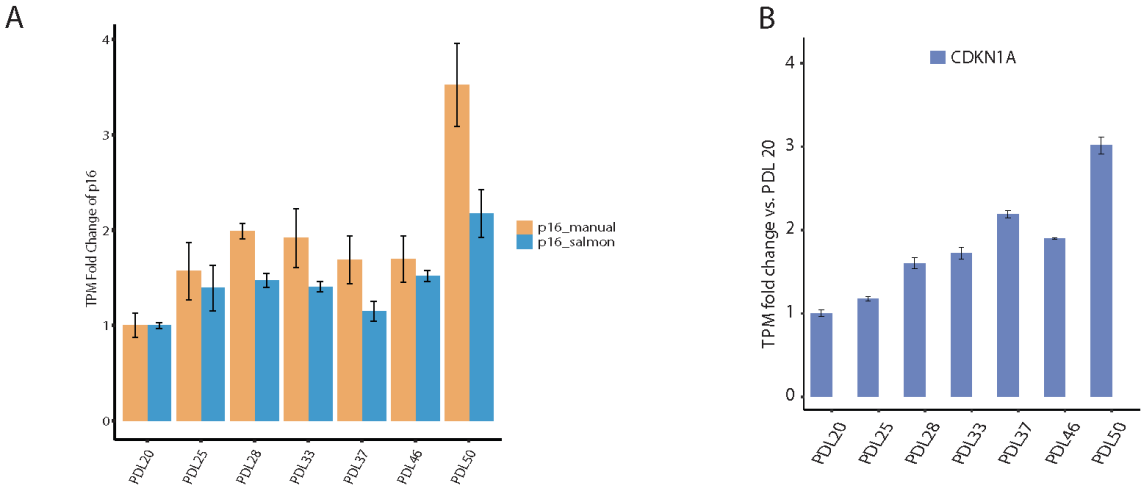

Figure S1: **Senescent WI-38 cells express classic senescence markers.** **(A)** Barplot of the expression of CDKN1A (p21) in WI-38 cells through the time course. Expression values are fold changes of TPMs normalized to PDL 20. Error bars represent the standard deviation of triplicate samples. **(B)**Barplot of expression levels of CDKN2A (p16) in WI-38 cells. (blue) TPM values generated by Salmon using the gencode v29 annotation set that includes p16 and p14 in one locus. (orange) a manual summation of the p16 only transcripts in the locus are shown. Error bars represent the standard deviation of three samples.

bioRxiv preprint doi: https://doi.org/10.1101/2021.05.03.442497; this version posted May 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

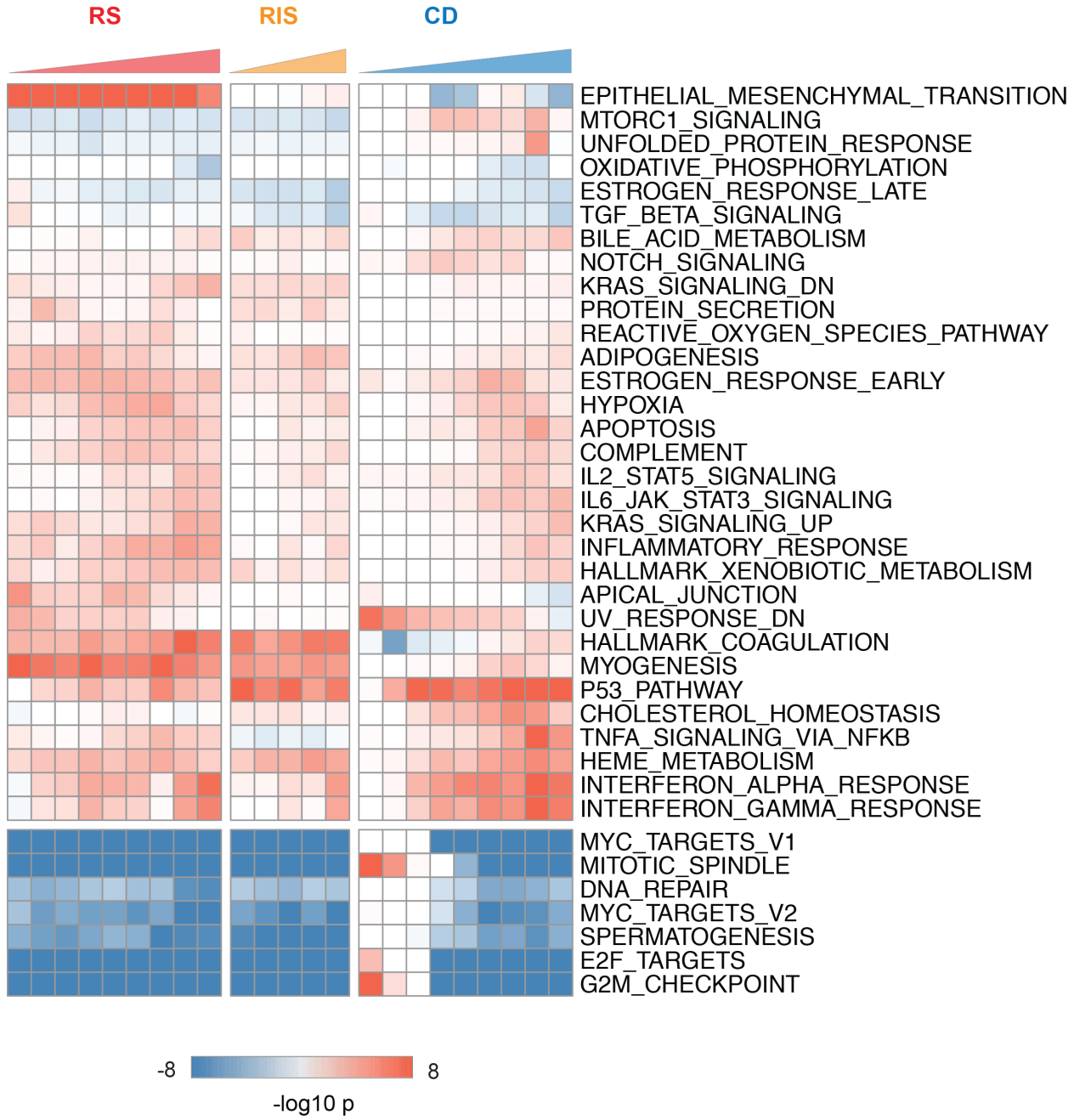

Figure S2: **Individual time point GSEA enrichments for replicative senescence (RS), radiation induced senescence (RIS), increasing cell density (CD). (A)** GSEA was run using misigDB Hallmark gene sets as the annotations. Data used for ranked list was the log2 fold changes at each time point vs. first the initial time point for each condiion. Color indicates -log10 p-value and direction of change e.g. red = enriched in up-regulated genes, blue=enriched in down-regulated genes

bioRxiv preprint doi: https://doi.org/10.1101/2021.05.03.442497; this version posted May 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

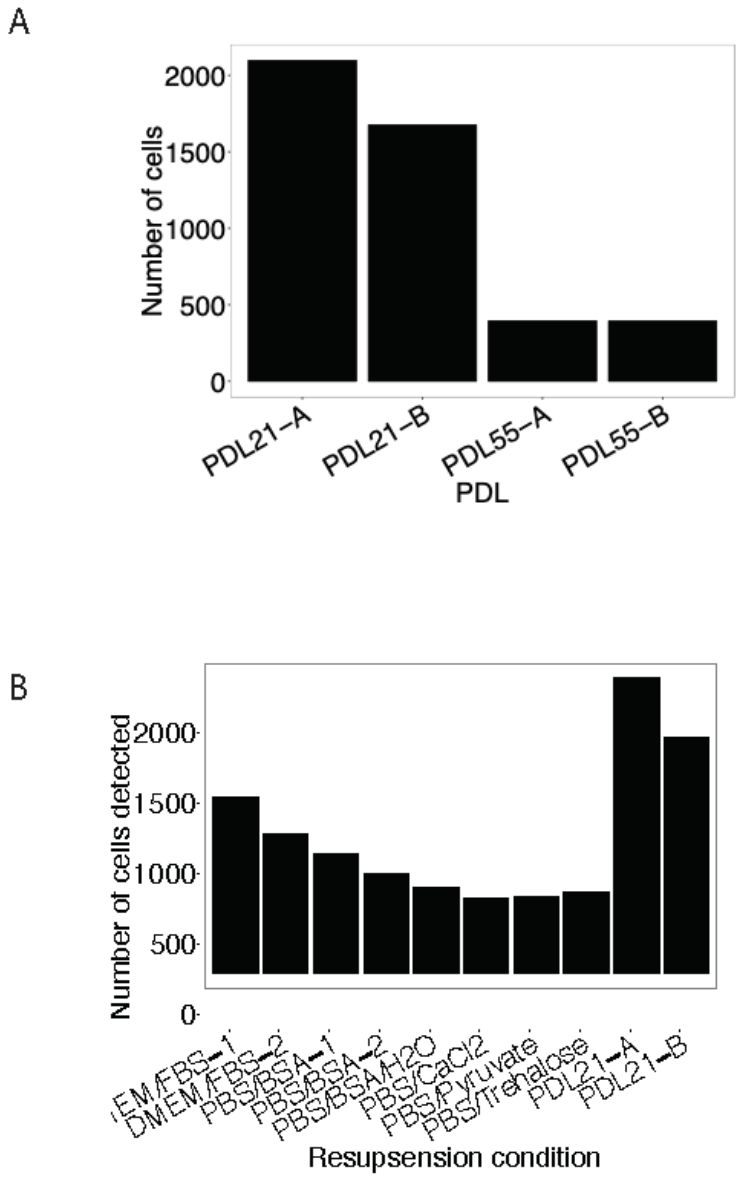

Figure S3: **Using a modified resuspension buffer of DMEM+FBS increases detection of senescent cells with the 10x Genomics 3' single cell RNA-seq protocol. (A)** Barplot comparing the number of cells detected (y-axis) as a function of the age of cells (x-axis). **(B)** Barplot representing the number of cells detected (y-axis) with various cell resuspension buffers (y-axis).

bioRxiv preprint doi: https://doi.org/10.1101/2021.05.03.442497; this version posted May 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

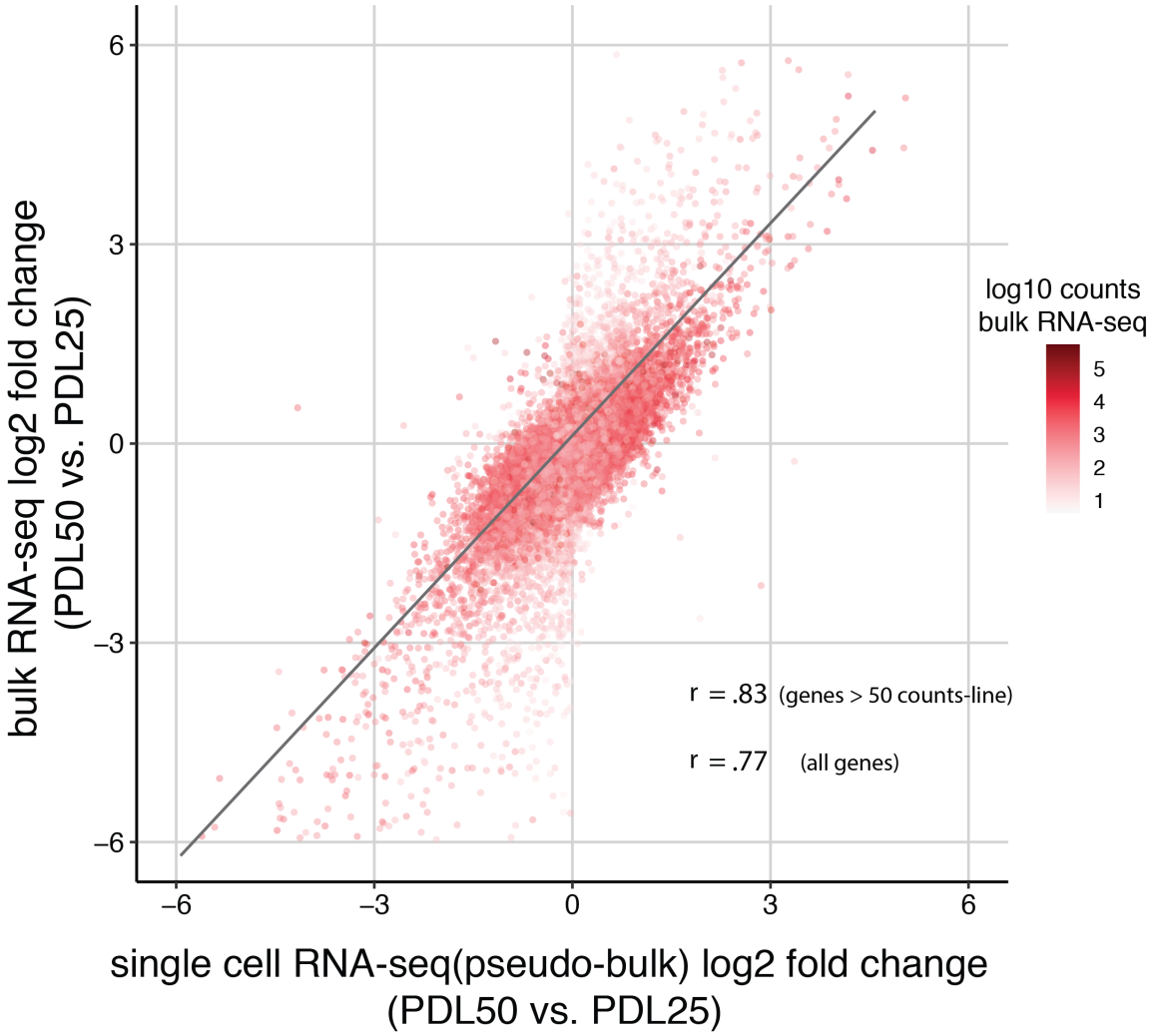

Figure S4: **RS dependent gene expression changes measured by Bulk RNA-seq and single cell RNA-seq are highly concordant. (A)** Scatter plot comparing the log2 fold changes of PDL50 vs PDL25 measured using single cell pseudo-bulk count summations (x-axis) and actual bulk RNA-seq (y-axis). The color for each gene is the log10 normalized counts. The r for all genes is 0.77. The r for genes filtered at > 50 counts is 0.83. The linear fit line was drawn using genes > 50 counts.

bioRxiv preprint doi: https://doi.org/10.1101/2021.05.03.442497; this version posted May 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

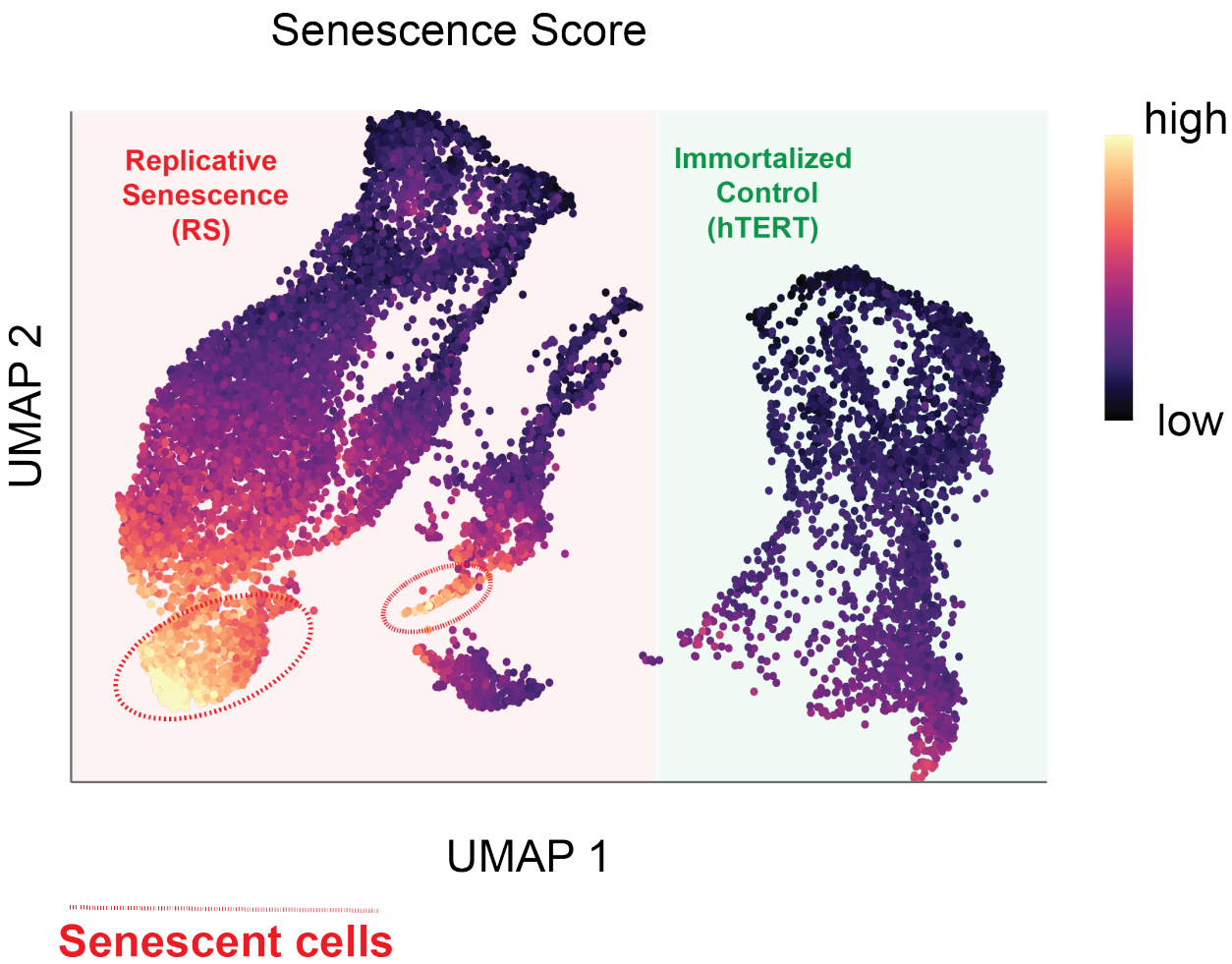

Figure S5: **UMAP projection of RS time course and hTERT cells colored with "senescence score" derived from bulk RNA-seq signature(top 1000 genes)**

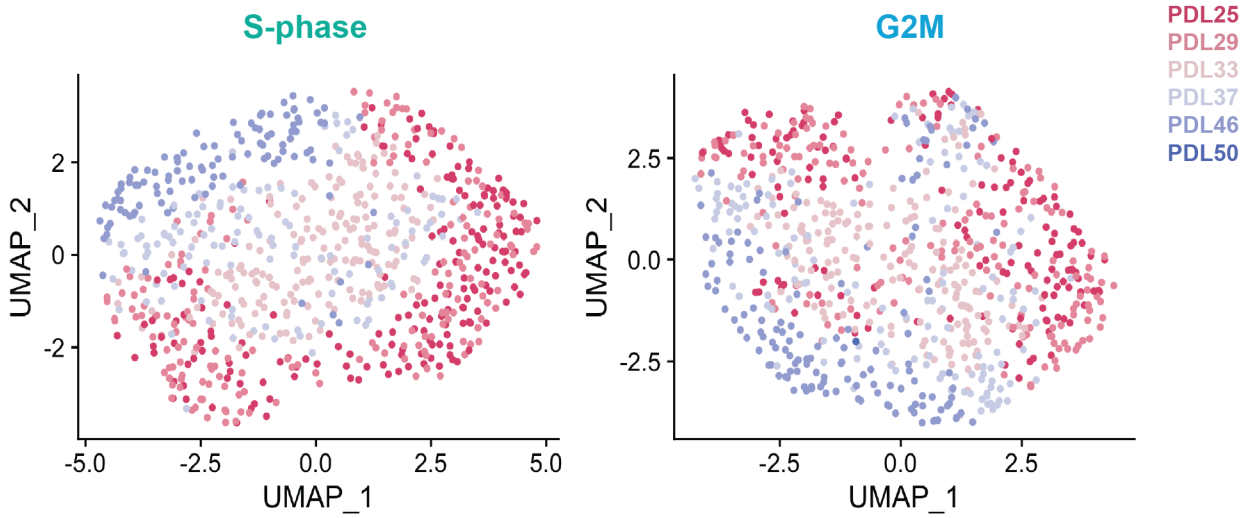

Figure S6: **Single cell UMAPs of mitotic Cells. (A)** S phase and G2M cells were isolated and reprocessed visualized with a UMAP projection ahowing that in mitotic cells, PDL is the primary source of variance across all PDLs

bioRxiv preprint doi: https://doi.org/10.1101/2021.05.03.442497; this version posted May 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

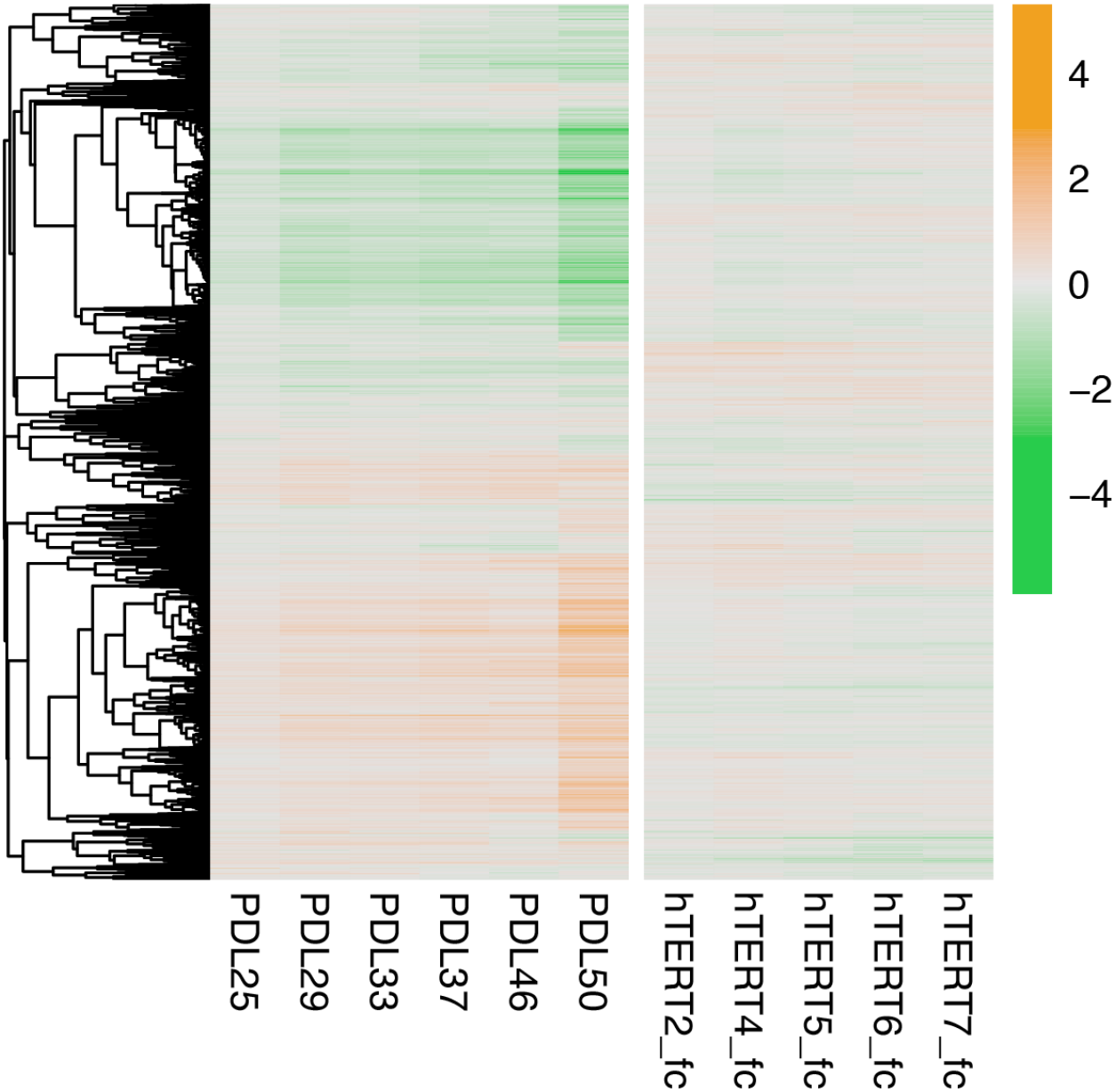

Figure S7: **PDL-dependent changes in the RS proteome vs hTERT cells (A)** hierarchical clustering of protein log2 fold changes at each time point/PDL versus first (not shown) for RS WT WI38 (left) and hTERT WI38 cells right from high induction (orange) to depletion (green)

bioRxiv preprint doi: https://doi.org/10.1101/2021.05.03.442497; this version posted May 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

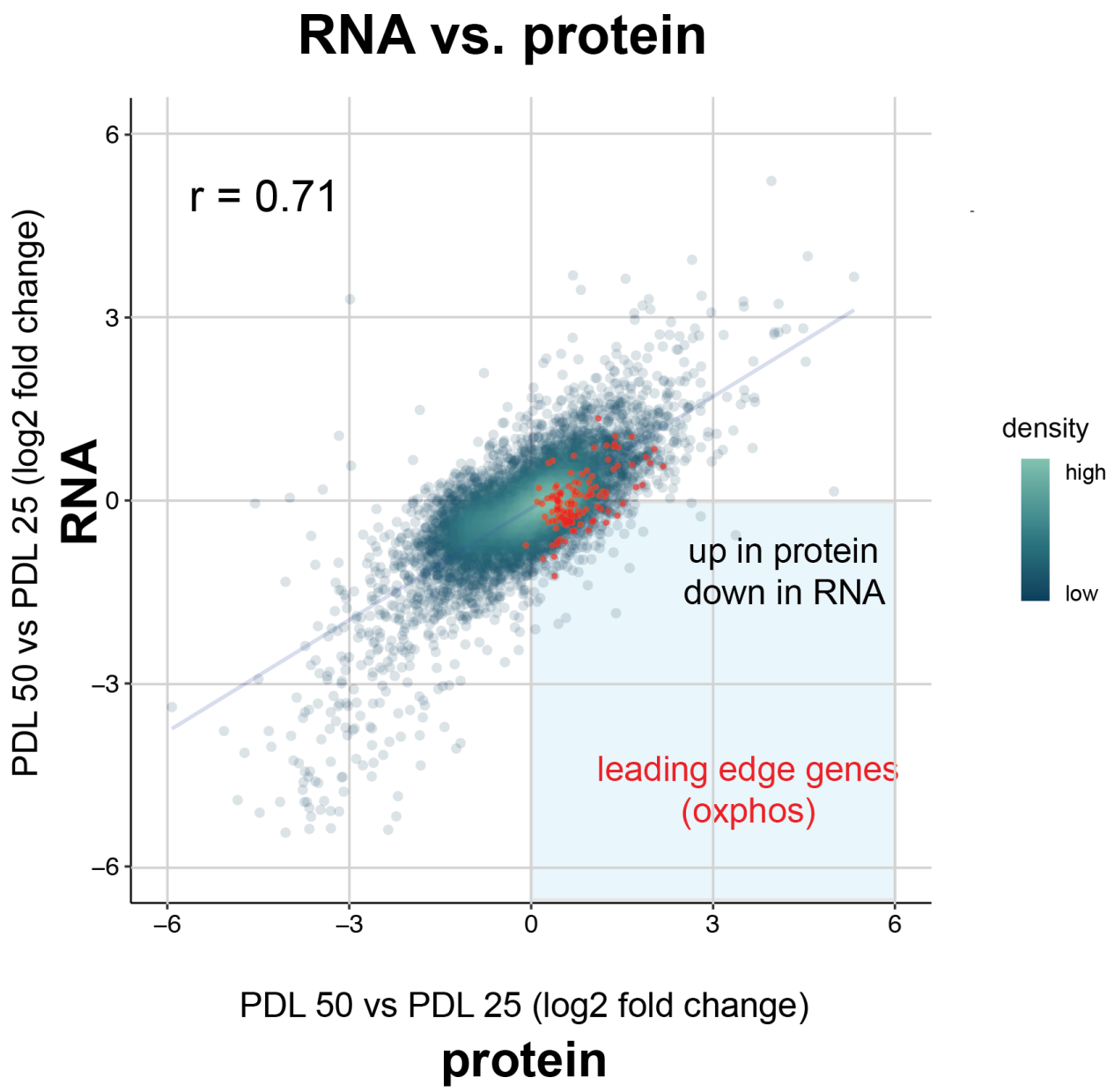

Figure S8: **Replicative senescence driven changes in the transcriptome and proteome are highly correlated (A)** Scatterplot comparing the log2 fold change expression of genes (x-axis) vs. protein (y-axis) in PDL 50 cells relative to PDL 20 cells. Oxphos leading edge genes from GSEA plotted in red. Point density colored from high (white) to low (green

bioRxiv preprint doi: https://doi.org/10.1101/2021.05.03.442497; this version posted May 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Figure S9: **Breakdown of the hallmark oxidative phosphorylation gene set into functional subsets reveals up regulation of most mitochondrial functions during RS (A)**Heatmaps of log2 FC proteomics data for the MsigDB hallmarks Oxidative Phosphorylation annotations set broken into constituent functional groups

bioRxiv preprint doi: https://doi.org/10.1101/2021.05.03.442497; this version posted May 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

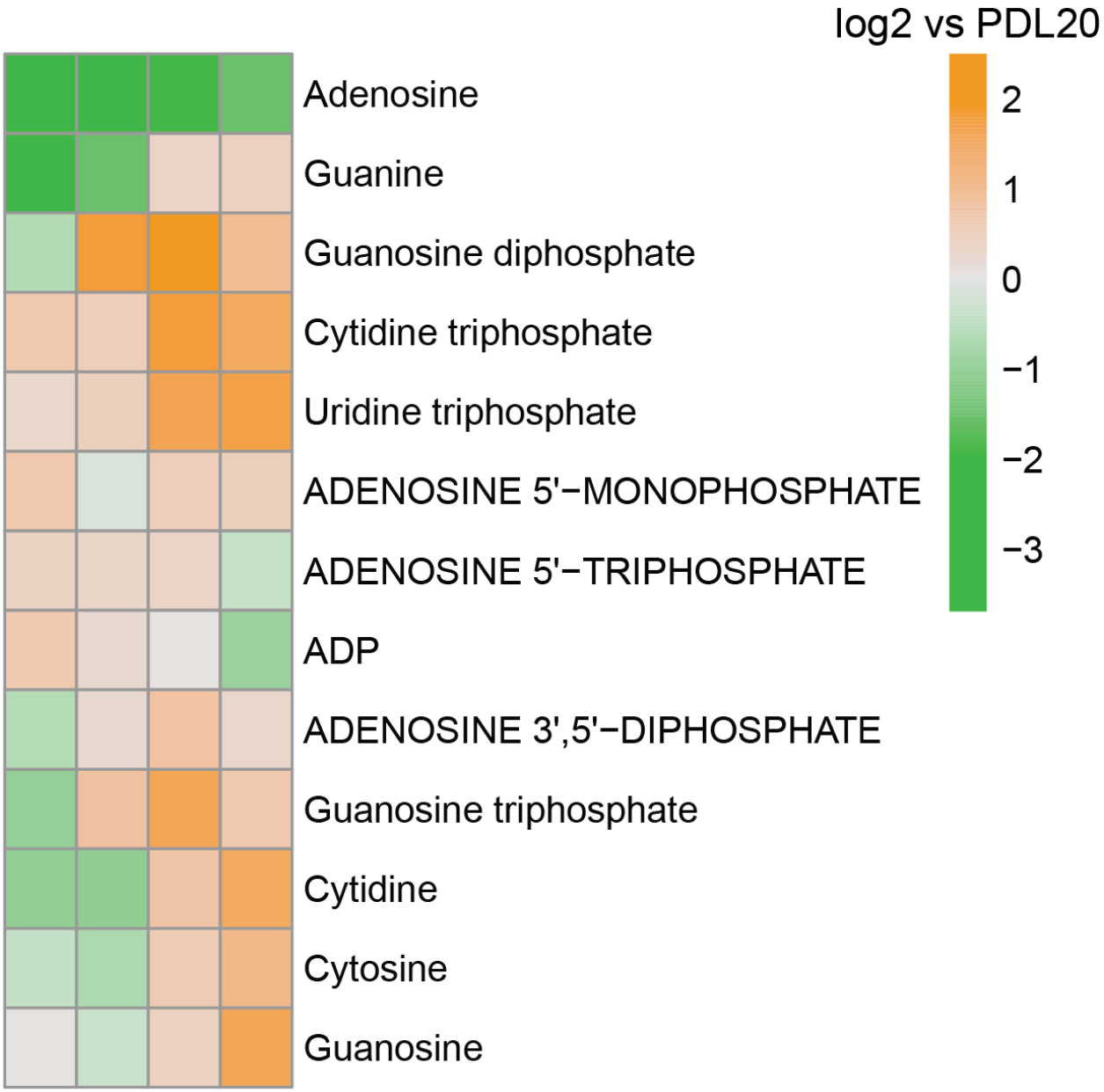

Figure S10: **Heatmap of log2 FC metabolomics data for nucleotides**

bioRxiv preprint doi: https://doi.org/10.1101/2021.05.03.442497; this version posted May 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

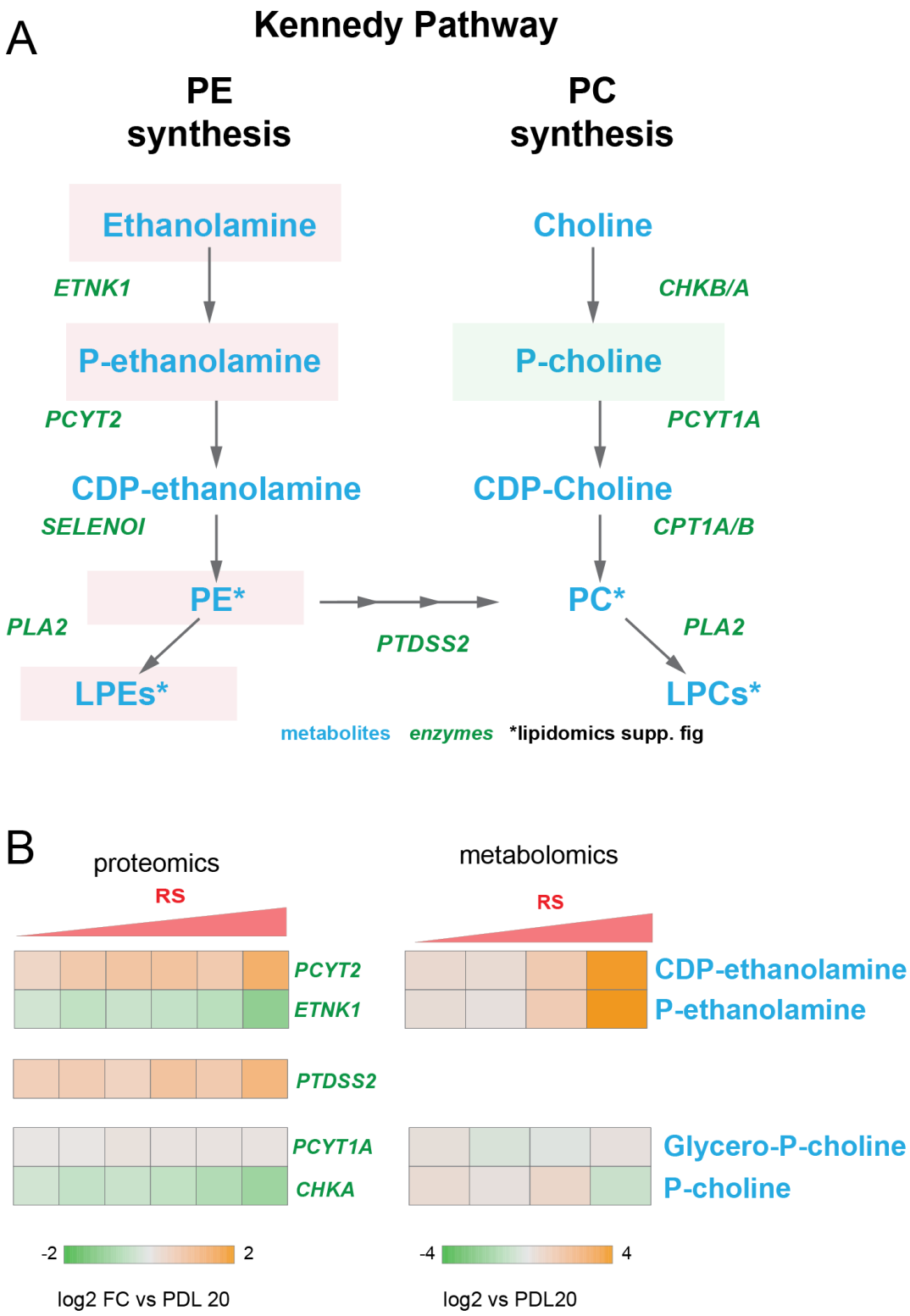

Figure S11: **Kennedy pathway utilization during RS (A)** Kennedy Pathway diagram. Metabolites in blue, proteins in italicized green **(B)** Heatmaps of log2 fold changes for metabolites and proteins from A

bioRxiv preprint doi: https://doi.org/10.1101/2021.05.03.442497; this version posted May 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

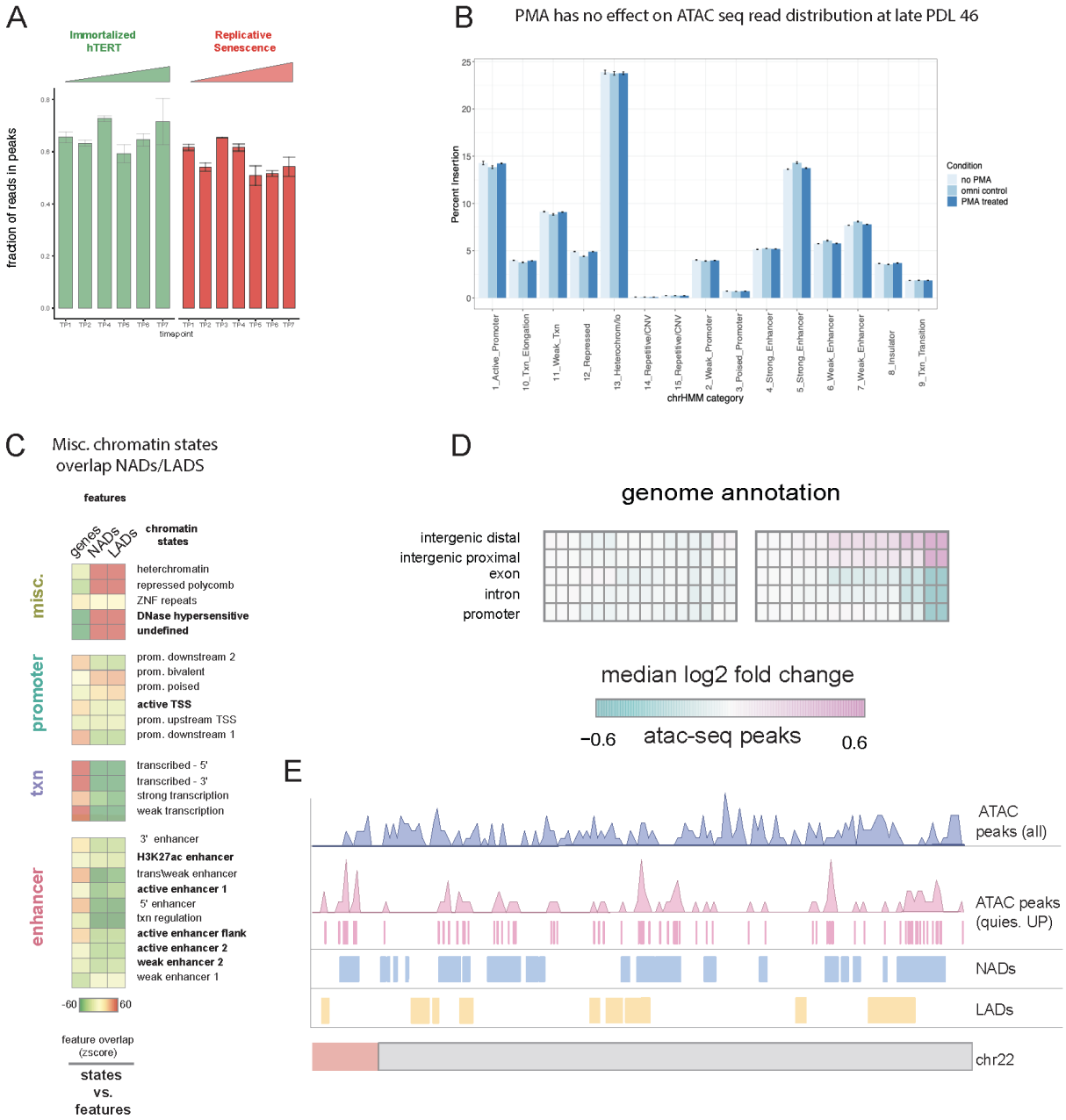

Figure S12: **RS ATAC-seq QC metrics, controls, and NADs/LADs browser shot. (A)** Fraction of reads in peak by PDL. **(B)** PDL 46 read distribution across chromatin states +/- PMA. **(C)** Chromatin state overlap Z-scores with genes/NADs and LADs. **(D)** Log2 fold change of peak accessibility in gene annotations. **(E)** Browser shot of ATAC peaks, ATAC peaks up in quiescent state, over NADs and LADs.

bioRxiv preprint doi: https://doi.org/10.1101/2021.05.03.442497; this version posted May 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

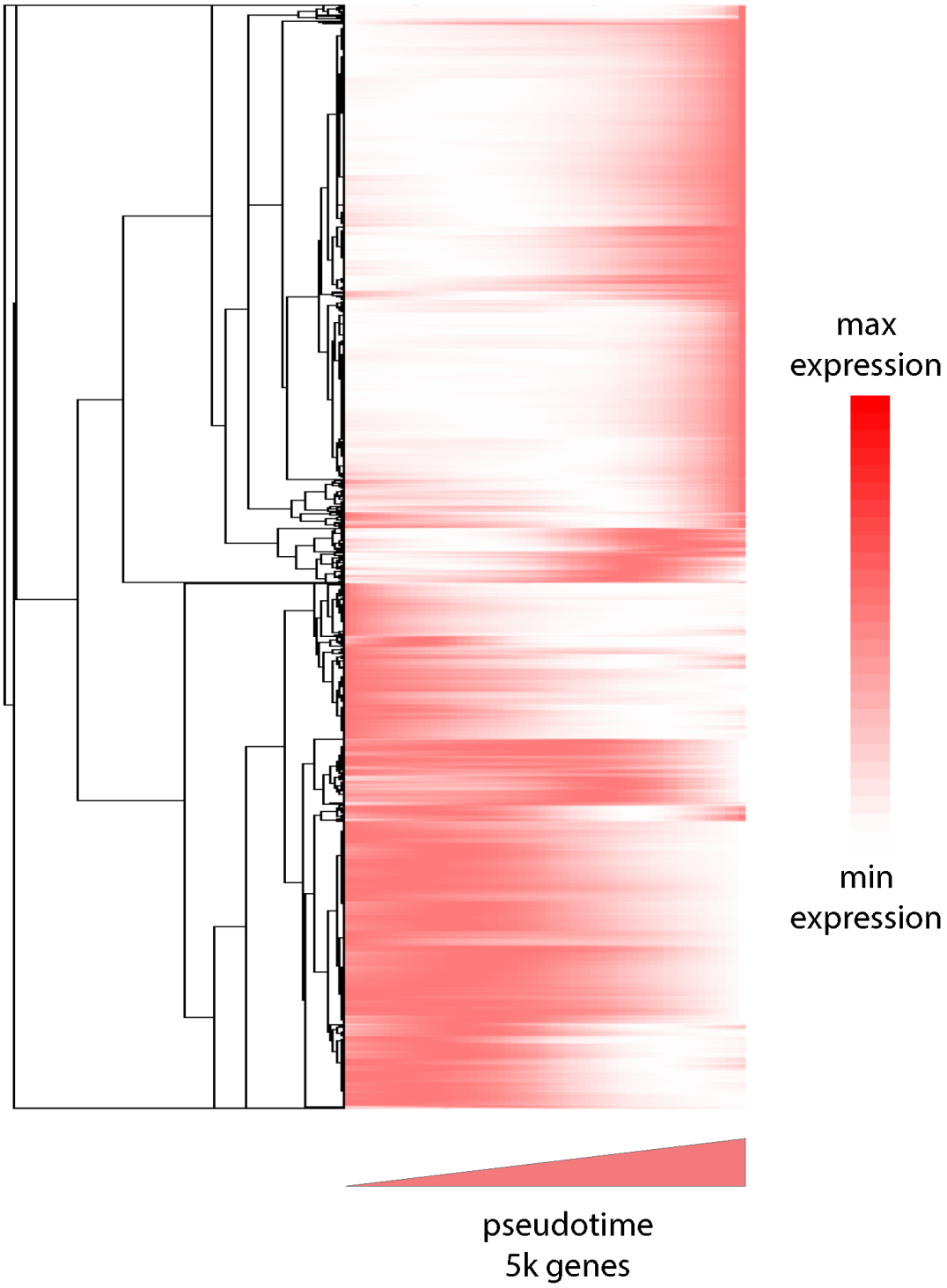

Figure S13: **Gene expression trajectories for the top 5,000 differentially expressed genes using Monocle3. (A)** Heatmap of the pseudotime trajectories for the top 5,000 differentially expressed genes. Genes were hierarchically clustered and gene expression levels are depicted (white = low expression, red = high expression).

bioRxiv preprint doi: https://doi.org/10.1101/2021.05.03.442497; this version posted May 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

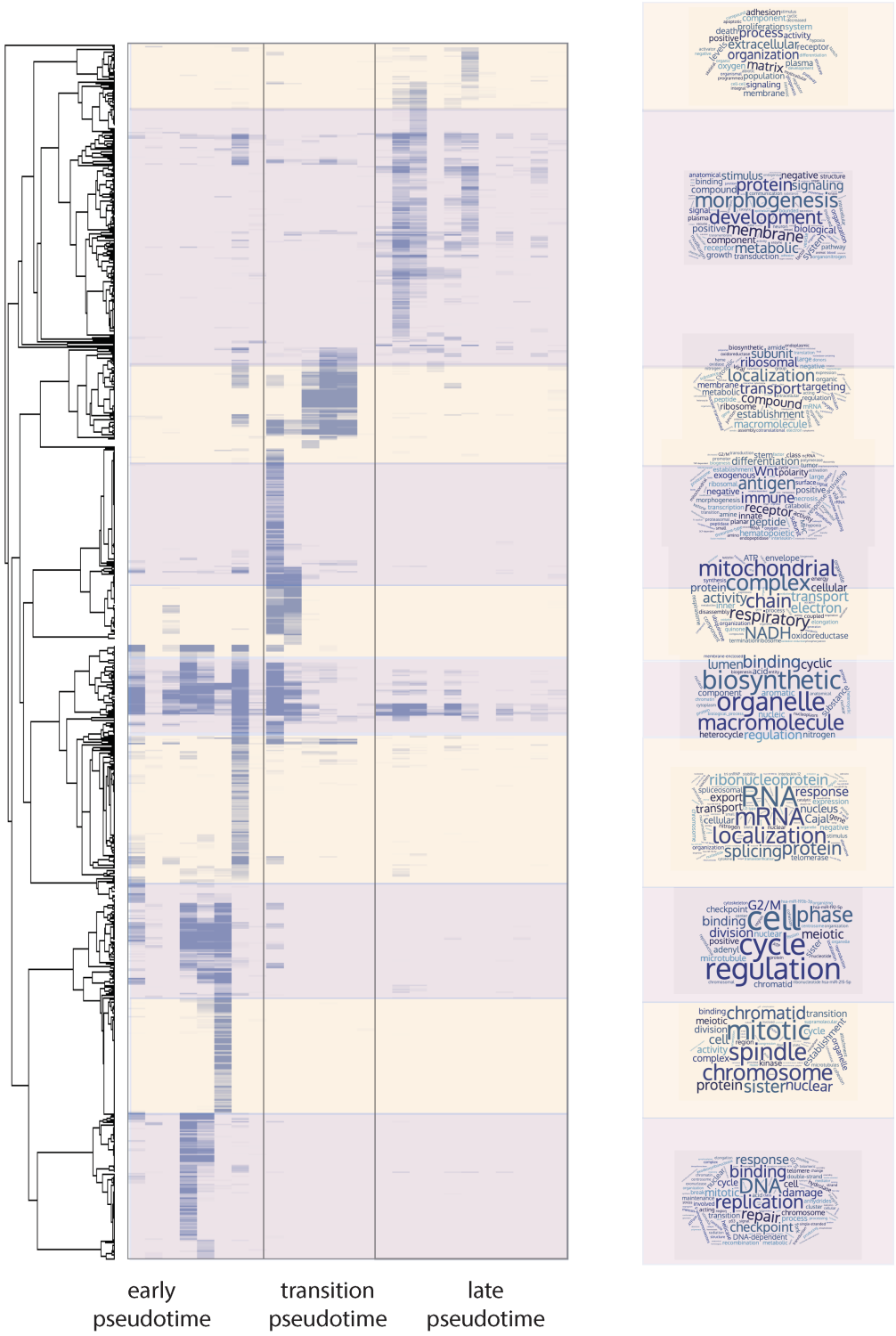

Figure S14: **Gene expression trajectories for the top 5,000 differentially expressed genes using Monocle3. (A)** Heatmap of p values for GO-term enrichment analysis on individual clusters from 8. Values plotted are -log10 p-value; light purple to dark purple. The ordered pseudotime clusters are shown on the x-axis while hierarchically clustered gene sets are shown on the y-axis. Hierarchical clustering of gene set enrichment was used to divide gene sets into 10 major clusters. GO-term names were used to create word clouds reflective of the types and functions of enriched gene sets.

bioRxiv preprint doi: https://doi.org/10.1101/2021.05.03.442497; this version posted May 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

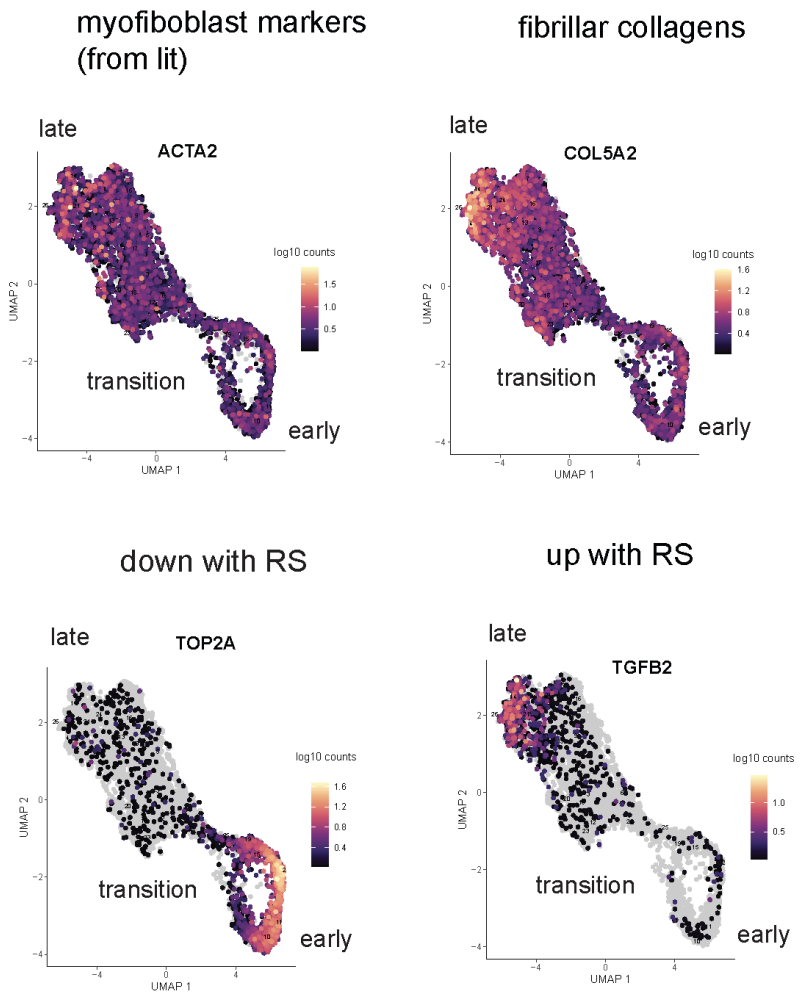

Figure S15: **UMAP examples of genes from Figure 8.**