Title

Pasta, a versatile transcriptomic clock, maps the chemical and genetic determinants of aging and rejuvenation

Authors

Jérôme Salignon[1,2#], Maria Tsiokou[3&], Patricia Marqués[3&], Enriqueta Rodríguez-Diaz[3], Hazel Ang[1,†], Federico Pietrocola[3#], Christian G. Riedel[1,2#]

1. Department of Medicine Huddinge, Karolinska Institute, Huddinge, Sweden
2. Department of Molecular Biosciences, The Wenner-Gren Institute, Stockholm University, Stockholm, Sweden
3. Department of Cell and Molecular Biology, Karolinska Institute, Solna, Sweden

[#]: Corresponding authors: jerome.salignon@ki.se, federico.pietrocola@ki.se, and christian.riedel@su.se

[&]: Equal contribution

† Present address: Institute of Molecular and Cell Biology, Agency of Science, Technology and Research, Singapore

Abstract

With the growing burden of age-related diseases, understanding and modulating the aging process has become a priority. Transcriptomic aging clocks (TACs) can track biological age but remain limited by platform dependence, tissue specificity, or restricted accessibility. To address this, we developed Pasta, a robust and broadly applicable human TAC built using a novel 'age-shift' learning framework. Pasta accurately predicted relative age across diverse tissues, data types, including bulk and single-cell RNA-Seq as well as microarray data, and species. Its predictions aligned with senescent and stem-like cellular states, with model coefficients enriched for p53 and DNA damage response pathways. Pasta's age scores correlated with tumor grade and patient survival in several cancer types, indicating potential clinical relevance. Applied to the Connectivity Map L1000 dataset, Pasta identified both established and previously unrecognized age-modulatory compounds and genetic perturbations, highlighting mitochondrial translation and mRNA splicing as key determinants of cellular propensity for aging and rejuvenation, respectively. Experimental validation confirmed pralatrexate as a potent senescence inducer and piperlongumine as a rejuvenating agent. Together, these findings establish Pasta as a versatile and accessible tool for aging research and therapeutic discovery.

Introduction

The growing prevalence of chronic non-communicable diseases in later life has positioned aging as a major driver of pathology. This, coupled with advances in understanding the fundamentals of aging and health[1,2], has intensified interest in pharmacological and lifestyle strategies to modulate aging[3]. Unlike chronological aging, which progresses at a constant pace, biological aging is modulated by endogenous and environmental factors that can

accelerate, decelerate, or even reverse aging processes[4]. This highlights biological age as a latent but measurable trait that reflects current and future health, captures individual variability in aging rates, and offers a means to evaluate geroprotective interventions.

Over the past decade, a range of computational tools—collectively termed 'aging clocks'—have been developed to quantify biological age and estimate aging rates. These models draw from diverse molecular and clinical hallmarks across multiple 'omics' layers, including epigenetic, transcriptomic, metabolomic, proteomic, and metagenomic data[5]. Among them, epigenetic clocks based on genome-wide CpG methylation patterns have been widely used due to their high precision in estimating chronological and biological age, strong predictive value for morbidity and mortality, and the relative stability of such DNA methylation (DNAm) marks, which enable the capture of cumulative aging effects[6]. However, DNAm clocks also have limitations, including reduced sensitivity to rapid or transient changes[7] and limited biological interpretability, as CpGs are not always clearly linked to gene function[8]. In addition, single-cell DNAm approaches remain costly, technically demanding, and lack standardized protocols, restricting their utility in resolving aging at single-cell resolution[6].

Transcriptomic aging clocks (TACs) address these limitations: they sense transient shifts, yield interpretable gene and pathway outputs, and can be applied to extensive bulk[9], single-cell[10,11], spatial[10,12], genetic[13,14] and chemical[13,15] perturbation datasets. Notably, ongoing efforts include constantly growing transcriptomic perturbation resources, such as the Connectivity Map (CMAP)L1000[13], Perturb-Seq datasets[14] or the Tahoe-100 dataset[15]. Applying TACs to such datasets could reveal regulators of aging and rejuvenation with translational potential. For example, compounds that raise the biological age of cancer cells may enhance therapeutic efficacy[16], and gene perturbations that accelerate aging in iPSC-derived neurons can improve models of late-onset disease[17]. Conversely, agents that lower age in normal cells may promote regeneration; improving stem cell fitness[18], iPSC protocols[19], or partial reprogramming cocktails[20]. While a handful of age-modulatory perturbations have recently been described[21], systematic approaches to identify them remain limited. A recent study showed the potential of this approach by mining the Gene Expression Omnibus (GEO)[22] database using aging clocks to find novel age-modulatory perturbations[23]. A landmark and untaped resource in this context is the CMAPL1000 dataset[13], which contains over three million transcriptomes from 248 cell lines exposed to more than 14,000 gene perturbations and 30,000 compounds. Identifying both general and cell line-specific age-modulatory perturbations in this dataset could inform the development of more effective age-related translational strategies.

Human TACs have been developed in blood[24], muscle[25], fibroblasts[26], skin[27], and multi-tissue settings[28–32], and more recently at single-cell resolution using blood[33], immune cells[34,35], and PBMCs[36]. Despite this progress, no TAC has achieved widespread adoption comparable to epigenetic clocks, due to limitations in performance, availability, and platform compatibility. Gene expression sensitivity to environmental cues enables detection of transient states but also introduces variability and noise, requiring robust modeling across heterogeneous datasets. Most TACs, including single-tissue models and the multi-tissue RNAAgeCalc[29], are trained on single datasets, leading to overfitting and limited generalization. Consistently, MultiTIMER[30], trained on multiple datasets, outperforms RNAAgeCalc[29] despite being trained on fewer samples (~3,000 vs ~9,600). Another key barrier is availability: DeepQA[32] and Shokhirev & Johnson[28]'s TACs offer no usable software, while RAPToR[31] requires users to generate custom references without default human datasets. Moreover, most multi-tissue TACs[28–30,32] require raw data reprocessing, batch correction, or model retraining, limiting their use to bulk RNA-Seq data and reducing reproducibility.

To address these limitations, we present Pasta (Predicting age-shift from transcriptomic analysis), a multi-platform, multi-tissue, ready-to-use TAC. It outperformed MultiTIMER[30], the currently best available multi-tissue human TAC, across diverse datasets, yielded meaningful predictions across tissues and platforms in mouse datasets, and had its predictions being largely influenced by p53-related genes. Beyond age prediction, Pasta effectively distinguished between senescent, quiescent, and stem cells, in both static and dynamic settings, indicating its capacity to capture bidirectional cell state transitions from stemness to senescence. This capability could be valuable in clinical contexts where Pasta may help to predict tumor grade. When applied to the CMAP L1000 dataset[13], Pasta highlighted potential age modulators such as pralatrexate and piperlongumine, which were experimentally validated, and pinpointed molecular features associated with cell-line responsiveness to pro- or anti-aging interventions. Taken together, Pasta offers a robust and interpretable framework for biological age estimation across tissues, platforms, and species, and for systematic discovery of genetic or chemical modulators of aging and rejuvenation, with potential relevance to translational research in oncology[16], neurodegeneration[17], and regenerative medicine[18–20].

Results

Construction of a general-purpose transcriptomic aging clock from heterogeneous multi-tissue data

Given that existing TACs often show limited transferability, being trained on single cohorts or requiring extensive preprocessing, we aimed to develop a broadly applicable transcriptomic aging clock that can be used across tissues and sequencing platforms. To ensure broad generalization, we compiled heterogeneous transcriptomic data from the Genotype-Tissue Expression (GTEx)[37] project, the GEO, and the Expression Atlas[38], retaining 17,212 healthy samples spanning multiple tissues from 21 studies (18 bulk RNA-Seq and 3 microarray; Fig. 1a, Supplementary Table 1). Gene expression values were rank-transformed to mitigate platform and processing effects.

Using this dataset, we compared three modeling strategies: regression models predicting chronological age (Fig. 1b), young-versus-old classifiers (Fig. 1c), and age-shift classifiers trained on sample pairs separated by a defined age gap (Fig. 1d) – the latter being an approach conceptually similar to twin neural network regression[39,40] and pairwise difference regression[41] or classification[42]. For the classifiers, predictions were converted to age scores (see Methods).

We evaluated all models using a leave-one-dataset-out (LODO) strategy. An age shift model with a 40-year age gap achieved the highest median Pearson correlation coefficient (PCC) across all samples (Fig. 1e) and remained the top performer when restricting the evaluation to young and old individuals only (Extended Data Fig. 1a). It also showed consistent performance across tissues and platforms (Extended Data Fig. 1b, 2, and 3). Models trained on unranked counts performed poorly (median PCC < 0.1), underscoring the importance of rank transformation (Fig. 1e; Extended Data Fig. 1a).

To assess generalization to data types unseen during training, we tested our top models on single-cell RNA-Seq datasets. Because training relied on bulk data, single-cell profiles were aggregated into pseudobulks of increasing size (1 to 1024 cells) to evaluate performance at different resolutions. Across retina, skin, and lung datasets, the age-shift model consistently outperformed the regression model and the young-versus-old classifier (Fig. 1f–h)[43–45].

Notably, the latter failed when test datasets lacked age categories present during training (Fig. 1f,h), highlighting the limitation of fixed age cutoffs.

Altogether, the 40-year age-shift model provided the most robust and generalizable predictions of relative age across heterogeneous transcriptomic data, outperforming previous approaches. We refer to this model as Pasta throughout the remainder of this study.

Pasta leverages p53-related genes to accurately predict relative age across platforms, tissues, and species

To evaluate Pasta's performance, we compared it with MultiTIMER, the most accurate available multi-tissue human TAC[30]. Pasta outperformed MultiTIMER in 7 of 8 datasets from our LODO analysis (Fig. 2a; Supplementary Table 2) and in 5 of 6 independent validation datasets (Fig. 2b). These datasets encompassed diverse tissues (Supplementary Table 2), confirming Pasta's broad applicability across tissue types. Since Pasta works across tissues and platforms, we wondered if it could also work across species. We extended Pasta for usage in mouse by leveraging one-to-one orthologues (see Methods). Encouragingly, we observed that Pasta could also predict age accurately in mouse across a variety of tissues and platforms (Fig. 2c,d). Altogether these results show that Pasta is an accurate predictor of relative age across tissues, platforms, and species.

We next conducted gene set enrichment analysis (GSEA) on broadly used collections (REACTOME, miRDB...)[46] using Pasta's coefficients. Genes with positive coefficients, whose higher expression predicted older age, were enriched for p53 signaling and irradiation response terms, pointing to activation of the DNA damage-p53 axis[47] (Fig. 2e; Supplementary Table 3). The second and third strongest negative enrichments, genes whose higher expression predicted younger age, were targets of miR29, a p53 activator[48] which drives aging phenotypes in mice[49]. Notably, no pathway was significantly enriched when using the baseline regression model generated in Figure 1, suggesting that the age-shift learning framework more effectively captures aging-relevant transcriptional signal.

Consistent patterns were observed at the individual gene level, with RANBP17 carrying the largest negative coefficient, nearly twice that of any other gene (Fig. 2f). This gene promotes cell proliferation[50] and has been proposed to be a master regulator of aging which decreases with age[51]. The strongest positive coefficients belonged to PCDHB3, a colorectal-cancer tumour-suppressor[52], CDKN2A (p16), a canonical p53 target and senescence marker, and ZMAT3, another key p53 target[53] (Fig. 2g).

In conclusion, Pasta is a robust multi-tissue, multi-platform, and multi-species transcriptomic age model which grounds its predictions in tumor-suppressor genes and core aging modulators.

Pasta accurately tracks cellular age across senescence, pluripotency and differentiation states

Given Pasta's strong performance across tissues, we then asked whether it could also track cellular aging in vitro, focusing on senescent and stem cells, which represent opposite ends of the cellular age spectrum.

Pasta accurately distinguished senescent from proliferative cells in 19 of 30 RNA-Seq datasets (AUC-ROC = 1) and achieved moderate discrimination (AUC-ROC > 0.7) in five additional datasets (Fig. 3a; Supplementary Table 4), consistently labeling senescent cells as

older. It also distinguished senescent from quiescent cells in 4 of 5 datasets (Fig. 3b). In a 450-day serial passaging experiment of primary fibroblasts that reached replicative senescence, Pasta's age scores increased almost linearly with culture duration (PCC = 0.896; Fig. 3c). We next assessed whether Pasta could detect chemical inducers of senescence in single-cell data. Combining two studies that tested the same 138 compounds in A549 cells, one using scRNA-Seq[54] and another using flow cytometry with a senescence reporter[55], Pasta assigned significantly higher age scores to senescence-inducing compounds than to other compounds ($p = 9.9 \times 10^{-6}$; Welch's two-sample t-test; Fig. 3d).

Having established that Pasta can detect senescent and thus very old cells, we next tested its sensitivity to very young cell states. Pasta effectively distinguished embryonic stem cells from fibroblasts across four bulk RNA-Seq datasets (Fig. 3e) and induced pluripotent stem cells from fibroblasts in five datasets (Fig. 3f), consistently labeling stem cells as younger. During OSKM-mediated fibroblast reprogramming, Pasta's age scores declined linearly with time (PCC = -0.905; Fig. 3g), whereas in directed differentiation of human embryonic stem cells toward liver cells, age scores increased steadily (PCC = 0.926; Fig. 3h).

Together, these results show that Pasta reliably tracks not only tissue age but also cellular age across a continuum of states, from senescence to pluripotency and differentiation.

Pasta reveals cancer type specific associations between cellular age, tumor grade, and patient survival

Studies using transcriptomic stemness models have shown that higher stemness correlates with advanced tumor grade and poorer survival in cancer patients[56,57]. Recently, increased cancer cell age has also been linked to adverse outcomes[16]. Indeed, the accumulation of senescent cells within tumors, whether normal or cancerous, can promote inflammation, relapse, and immune evasion[16,58–60]. Yet, data from The Cancer Genome Atlas (TCGA), the largest transcriptomic cancer resource, have not been systematically examined using an aging clock, leaving it unclear whether poor prognosis is driven by high stemness alone or also by senescence.

To test this, we applied Pasta to TCGA tumors, leveraging its capacity to identify both youthful (high-stemness) and aged (senescent-like) transcriptional states. Pasta's score showed weak correlation with donor age (PCC=0.0570), so unadjusted values were used in subsequent analyses. Six of twelve grade annotated cancer types showed significantly negative correlations between age scores and tumor grade (Fig. 3i, j, Supplementary Table 5), while two—kidney renal clear cell carcinoma (KIRC) and pancreatic adenocarcinoma (PAAD)—showed significantly positive correlations (Fig. 3i, k). These results indicate that both increased and decreased biological age can be associated with higher tumor grade, consistent with roles of both stemness and senescence in tumor progression. Donor age significantly predicted grade in only two cancers (Lower Grade Glioma, LGG, and Uterine Corpus Endometrial Carcinoma, UCEC). Pasta's stronger associations with tumor grade than donor age suggest that its scores capture biological rather than chronological age.

To test whether Pasta's age scores were also linked to clinical outcomes, we fitted Cox proportional hazards models across 33 cancer types. After multiple testing correction, younger age scores were significantly associated with poorer survival in Adrenocortical carcinoma (ACC), Sarcoma (SARC), and Skin Cutaneous Melanoma (SKCM) (Fig. 3l,m; Supplementary Table 6), whereas older age scores were significantly associated with poorer survival in Thymoma (THYM) and Acute Myeloid Leukemia (LAML) (Fig. 3l,n). Several

other cancers showed strong associations. Chronological age was a weaker predictor of survival in 18 cancer types and was uniformly associated with poorer outcomes.

Taken together, Pasta outperformed chronological age as a predictor of tumor grade and survival in several cancer types.

Systematic identification of age-increasing and rejuvenating small molecules in the CMAP L1000 dataset

Having established Pasta's predictive power for tissue and cellular aging in health and disease, we next assessed whether it can also identify age-modulatory perturbations.

To this end, we applied Pasta to the CMAP L1000 dataset[13], which contained the transcriptomic responses of 248 cell lines to more than 30,000 chemical and 14,000 genetic perturbations, totaling more than three million transcriptomes. Coincidentally, the bead-based nature of the L1000 assay provided also an opportunity to test Pasta's cross-platform robustness. Age scores were calculated for each perturbation and plate-normalized to corresponding negative controls (DMSO for compounds, empty vectors for genetic perturbations).

By this approach, we identified 259 compounds that were significantly associated with higher age scores and 59 compounds with lower age scores, hereafter termed *Aging* and *Rejuvenating* compounds, respectively (Fig. 4a; Supplementary Table 7). As expected, top Aging compounds included canonical senescence inducers such as mitoxantrone, gemcitabine, and doxorubicin[61], whereas top rejuvenating compounds included pluripotency-promoting molecules, such as RepSox and PD0325901[62]. Annotating the CMAP library with known senescence-[61] and reprogramming-[62] related drug classes revealed that 18.9% of Aging compounds belonged to senescence-inducing classes, compared with less than 0.2% among all remaining compounds ($p = 5.45 \times 10^{-76}$, Fisher's exact test, Fig. 4b). Similarly, 16.9% of Rejuvenating compounds belonged to reprogramming-related classes, compared with less than 0.2% among all remaining compounds ($p = 3.01 \times 10^{-18}$, Fisher's exact test). Notably, 27.1% of Rejuvenating compounds were annotated as both pro-senescence and pro-reprogramming, all of which were histone deacetylases (HDAC) inhibitors, a drug class known to induce both reprogramming[63,64] and senescence[65–67].

When repeating the analysis using the baseline regression model from Figure 1, we found the regression model to be inferior to Pasta, with the latter recovering significantly more senescence-inducing compounds (12.7% vs 18.9%; $p = 0.025$; Fisher's exact test; Extended Data Fig. 4). A similar trend was observed for the prediction of reprogramming-promoting compounds (6.5% vs 16.9%; $p = 0.142$; Fisher's exact test; Extended Data Fig. 4). Taken together, this data showed that Pasta identified age-modulatory compounds more effectively than a conventional regression model.

Analysis of the mechanisms of action (MOAs) annotated for the CMAP compounds revealed 26 enriched MOAs among Aging compounds, including inhibition of topoisomerases, cyclin-dependent kinases (CDKs), mouse double minute (MDMs), and inosine 5′-monophosphate dehydrogenase (IMPDHs), consistent with DNA damage and growth arrest pathways (Fig. 4c; Supplementary Table 8). Six MOAs were enriched among Rejuvenating compounds (Fig. 4d), dominated by HDAC inhibition, in line with their known role in accelerating reprogramming[63,64]. Additional Rejuvenating MOAs included MAPK/ERK kinase (MEK) and glycogen synthase kinase 3 (GSK3) inhibition, common components of reprogramming cocktails[62], as well as cardiac glycosides and proteasome inhibitors, both previously described

as senolytics[68]. Such senolytics may rejuvenate transcriptomic profiles indirectly, through selective elimination of aged cells (see Discussion).

Analyzing compound effects across cell lines revealed both conserved and context-dependent responses. Proteasome inhibitors MG-132 and bortezomib consistently reduced age scores, while mitoxantrone increased them (Fig. 4e; Supplementary Table 9). In contrast, several HDAC inhibitors (e.g., trichostatin-A, vorinostat, and panobinostat) showed divergent effects, reducing age scores in most cell lines but increasing them in neuronal lines (NEU, NPC, SH-SY5Y), consistent with prior reports of trichostatin-A's cytotoxicity in SH-SY5Y cells[69]. These results underscore the dual and context-dependent roles of HDAC inhibitors in promoting both reprogramming and senescence[63–67].

Chemotherapeutics have been described as having pro-aging properties[56]. Interestingly, we observed that most approved anti-cancer drugs were enriched in our set of Aging compounds (Supplementary Note 1, Supplementary Tables 10-12, Extended Data Fig. 5). Additionally, we identified compounds that had a stronger pro-aging effect in cancer cell lines than in normal cell lines. Such candidates might have translational applications for more effective and lower toxicity anti-cancer therapies. Furthermore, we could show that increase in cellular age is most often decoupled from increase in cell death (Supplementary Note 2, Supplementary Tables 13-15, Extended Data Fig. 6 and 7). This indicates that applying aging clocks to perturbation datasets could lead to novel anti-cancer therapies that have evaded conventional cell viability screens.

Collectively, these results demonstrated Pasta's ability to detect both aging- and rejuvenation-inducing compounds in the CMAP L1000 data, underscoring its cross-platform robustness and ability to reveal conserved and context-dependent age-regulatory mechanisms.

Experimental validation of compounds with age-modulatory effects

To experimentally test Pasta's predictive power in identifying cell-type-specific age-modulatory compounds, we selected two top-ranked hits, one predicted to increase and one to decrease cellular age. For each compound, two cancer cell lines were chosen: a predicted *responder* expected to show the age-modulatory phenotype and a *non-responder* predicted to remain largely unaffected (Supplementary Table 16).

We first examined pralatrexate, the fourth-ranked Aging compound among 30,920 in the CMAP dataset (Supplementary Table 7). In the responder A375 melanoma cells, pralatrexate induced clear signs of senescence, including cell enlargement and positivity for the Senescence-Associated β-Galactosidase assay (Fig. 5a). It increased CDKN1A (p21) and IL6 (SASP component) expression (Fig. 5b), reduced proliferation as indicated by lower EdU incorporation (Fig. 5c), and downregulated LMNB1 while upregulating p21 protein (Fig. 5d), all consistent with a *bona fide* senescence response. In contrast, the non-responder MDA-MB-231 breast cancer cells displayed only mild stress responses lacking key hallmarks of cellular senescence: minimal morphological changes yet detectable SA-β-Gal positivity (Fig. 5e), negligible changes in CDKN1A and IL6 mRNA levels (Fig. 5f), increased proliferation and EdU incorporation (Fig. 5g), and modest LMNB1 variation accompanied by upregulation of p21 protein levels (Fig. 5h).

Next, we evaluated Piperlongumine, ranked twentieth among 30,920 rejuvenating compounds (Supplementary Table 7). In the responder PC3 prostate cancer cells, piperlongumine enhanced expression of the pluripotency factors OCT4, SOX2, and NANOG and increased Oct4 protein abundance (Fig. 5i,j), indicating a transcriptional shift towards a youthful, stem-

like state. In contrast, non-responder MCF7 breast cancer cells showed none of these phenotypes (Fig. 5k,l).

Together, these findings confirm that Pasta accurately predicts cell-type-specific age-modulatory outcomes, validating pralatrexate as a senescence-inducing and piperlongumine as a stemness-promoting compound.

Systematic identification of age-increasing and rejuvenating genetic perturbations in the CMAP L1000 dataset

To uncover genetic drivers of cellular aging, we analyzed gene knockdowns, knockouts, and overexpression profiles in the CMAP L1000 dataset. Across cell lines, 841 gene perturbations (GPs) were significantly associated with higher age scores (*Aging GPs*) and 54 with lower age scores (*Rejuvenating GPs*).

The top two Aging GPs were overexpression of KRAS and BRAF (Fig. 6a; Supplementary Table 17). Furthermore, overexpression of three additional genes in the mitogen activated protein kinase (MAPK) pathway (RIT1, SOS1, and MAPK1) were in the top 12 pro-aging GPs. This result is consistent with the tumor-suppressive mechanism of oncogene-induced senescence (OIS), which is commonly triggered by hyperactivation of the MAPK pathway[70]. The third best Aging GP was knockout of CCNA2, consistent with a prior study reporting potent senescence induction by CCNA2 depletion[71]. The fourth and fifth best Aging GPs were knockout of the cyclin CCNA2 and knockdowns of ERG and MYC. Interestingly, even though ERG and MYC are also oncogenes, their overexpression does not seem to trigger OIS but instead repress it[72,73]. Accordingly, MYC depletion has been shown to promote senescence[73,74].

The top Rejuvenating GP was again connected to MYC – namely its overexpression – consistent with its key role in promoting stemness[75] and repressing OIS[73] (Fig. 6a; Supplementary Table 17). The second best Rejuvenating GP was SH2B3 knockdown, which aligns with the reported improvement in hematopoietic stem cell function upon Sh2b3 knockdown in *Fancd2*$^{-/-}$ mice[76]. The fourth best Rejuvenating GP was SMAD4 knockdown, which is consistent with SMAD4's role in repressing self-renewal of neuroblastoma stem cells[77]. Finally, TP53 knockout and knockdown ranked as the 12th and 18th best Rejuvenating GPs. This result is consistent with the central tumor-suppressive role of TP53 and its mutation in 50-60% of human cancers[78].

We then systematically characterized the significant age-modulatory GPs beyond the top hits. To this end, we inverted overexpression effects to pool all GP results and conducted gene set enrichment analysis using two complementary databases. The Aging Hallmark gene sets from the Open Genes database[79] were used to explore aging signatures, and the REACTOME database[80] gene sets were used to comprehensively study detailed cellular pathways (Fig. 6b–f; Extended Data Fig. 8a-f; Supplementary Tables 18 and 19). Among Rejuvenating GPs, the only enriched Aging Hallmark[79] term was "stem cells" (Fig. 6b), while among Aging GPs the most enriched term was "senescent cells" (Fig. 6c). Aging GPs were also enriched for transcriptional regulation, insulin/IGF-1, mTOR, and telomeres pathways, all of which are related to aging and its regulation. In REACTOME, Rejuvenating GPs were enriched in RAF/MEK/ERK and FGFR signaling (Fig. 6d), whereas Aging GPs were enriched in pathways linked to cell cycle progression, receptor tyrosine kinase activation (KIT, VEGF, NTRK), PI3K/AKT signaling, and immune activation (TLR4, CD28) (Fig. 6e). To get a consolidated list of age-modulatory perturbations, we identified genes for which both genetic perturbations and compounds targeting them produced significant age effects in the same

direction (Extended Data Fig. 8g). These genes were enriched in p53-dependent cell-cycle control and growth-factor signaling (VEGF, ERBB2, KIT, extranuclear estrogen) (Fig. 6f).

Next, we identified GPs with strong cell-line-specific age effects. MYC overexpression repeatedly produced rejuvenation, while MYC knockdown, CCNA2 knockout, and BRAF/KRAS overexpression increased cellular age (Fig. 6g; Supplementary Table 20). Others, such as SH2B3 knockdown and ASAH1 knockout, could either increase or decrease cellular age depending on context, underscoring the cell-type specificity of age regulation.

In summary, these findings map the genetic architecture of transcriptomic age modulation, revealing broad regulators like BRAF and MYC as well as context-dependent determinants of cellular age.

Defining the molecular determinants of the propensity of cells to respond to age-modulatory perturbations

While we previously revealed the pathways that drive transcriptomic age prediction (Fig. 2) and that induce age-modulation (Fig. 4), it remains unclear which pathways determine the heterogeneity in responses of various cell lines to the same pro-aging or rejuvenating cues. We addressed this question by leveraging the extensive CMAP L1000 and DepMap datasets, by identifying molecular features that govern a cell's propensity to age or rejuvenate. *Aging and rejuvenation propensities* were computed for 19 cell lines by averaging cell line-specific age scores of globally significant age-modulatory perturbations (see Methods). This analysis revealed marked heterogeneity (Fig. 7a), with MDA-MB-231 showing the highest rejuvenation propensity (-14.2 vs. -9.6 in MCF10A), consistent with its role as a model for aggressive, triple-negative breast cancer[81].

When aging propensity was correlated with the number of non-silent mutations, TP53 emerged as the only significantly associated gene (PCC = 0.875; Fig. 7b; Supplementary Table 21). Cell lines harboring two TP53 mutations (i.e., JURKAT) had negative aging-propensity scores, those with one mutation scored below six, and TP53 wild-type lines scored higher (Fig. 7c). These results align with TP53's established role as a tumor suppressor that enforces senescence in highly proliferative cells[82,83] and suggest novel potential tumor suppressor genes (Fig. 7b). Consistent with this idea, top significant genes in Fig. 7b included the known tumor suppressor genes FAT3[84], HERC1[85], and ABR[86,87].

To identify pathways influencing aging and rejuvenation propensities, we performed REACTOME enrichment analysis on proteins, transcripts, and gene copy numbers correlated with these traits ($|PCC| > 0.4$, Fig. 7d–i; Extended Data Fig. 9; Supplementary Table 22). Proteins and transcripts associated with aging propensity were both enriched in mitochondrial translation (Fig. 7d,e), underscoring mitochondrial activity as a central determinant of age-increase. This echoes recent findings linking p53, mitochondrial translation, and the SASP (see Discussion)[88]. Copy number–based enrichments included interferon-related pathways (Fig. 8f), consistent with type I interferons' role in SASP[89], DNA damage–induced[90] and oncogene-induced senescence[91], as well as their pro-aging effects[92].

Rejuvenation propensity correlated with proteins involved in mRNA splicing, histone acetylation, and chromatin modification (Fig. 7g). At the copy number level, HDAC and PRC2 pathways showed the strongest negative associations (Fig. 7h). These findings align with evidence that HDAC inhibition promotes reprogramming[63,64] and that PRC2 supports pluripotency[93] and functions as a robust biomarker of aging and rejuvenation[94–96]. Proteins

negatively correlated with rejuvenation propensity were enriched in apoptosis and checkpoint modules (Fig. 7i). This result is consistent with the finding that elevated expression of the Cdkn2a/Arf tumor-suppressor locus, which regulates senescence and apoptosis, constitutes a barrier for cellular reprogramming[97].

Together, these results delineate the molecular architecture of cellular age control, where mitochondrial translation and chromatin regulation act as opposing axes.

Discussion

In this study, we introduce Pasta, a versatile multi-tissue transcriptomic aging clock that robustly predicts relative age across diverse human and mouse transcriptomic datasets. Pasta outperformed existing transcriptomic clocks, captured cellular senescence and stemness states, and identified genetic and chemical modulators of aging. It is available as a user-friendly, open-source R package to support broad use in aging research.

A key strength of Pasta lies in its adaptability to diverse transcriptomic platforms, which was achieved by using an age-shift learning framework trained on heterogeneously processed RNA-Seq and microarray datasets. This design enhanced cross-dataset robustness and allows users to input gene-expression count data directly, without cross-dataset normalization or batch correction. Pasta achieved strong accuracy across tissues and data types and relied on biologically relevant features that produced more meaningful predictions than conventional regression models, positioning age-shift learning as an effective approach for predicting relative age.

Pasta's gene coefficients and enrichment analyses converge on the DNA damage response and its downstream effector, p53 signaling, as key pathways linked to transcriptomic aging. Core predictors such as CDKN2A/p16 and ZMAT3, both canonical p53 targets, underscore this mechanistic grounding. Downstream perturbation analyses reinforced these links: topoisomerase inhibitors, which induce DNA damage, and CDK inhibitors, which block the cell cycle, were the two most enriched mechanisms of action among aging-associated compounds, and pathways related to G1/S transition and TP53 regulation were consistently enriched across chemical and genetic pro-aging perturbations. Notably, recent work has proposed the DNA damage response as a fundamental "counting unit" of aging clocks, in which DDR, cell-cycle progression, and senescence act as interconnected temporal layers of cellular timekeeping[98]. Within this framework, the transcriptomic aging patterns captured by Pasta may reflect features emerging from these distinct but coordinated biological timescales.

Pasta showed strong predictive power for identifying cellular senescence, stemness, and age-related dynamics such as reprogramming and differentiation, despite not being trained on senescence or stem cell data. This suggests that transcriptomic aging clocks capture local cellular aging effects along a continuum from stemness to senescence. Future studies could assess whether aging clocks trained on other molecular layers, such as DNA methylation or proteomics, display similar behavior.

Pasta's age scores correlated significantly with tumor grade and survival across multiple cancer types, supporting its potential use for patient stratification. Moreover, Pasta's age scores showed a stronger association with tumor grade than donor age, suggesting that the model captures biological aging signatures beyond chronological age.

Analysis of the CMAP data showed marked enrichment of senescence-inducing perturbations in the Aging group and of reprogramming- and stemness-associated perturbations in the Rejuvenating group. Notably, Aging perturbations (259 compounds, 841 GPs) greatly outnumbered Rejuvenating ones (59 compounds, 54 GPs), consistent with a recent report[23], suggesting that diverse perturbations can accelerate aging, whereas rejuvenation requires higher precision. Perturbation effects were largely consistent across both cancer and normal (as defined in the CMAP cell annotation file) cell lines (Fig. 4e, 6g, Supplementary Tables 9, 16), indicating conservation of aging and rejuvenation pathways in health and disease. Interestingly, chemotherapeutics were strongly enriched in pro-aging compounds. These results are consistent with the view that many cancer drugs act by inducing senescence[99] and increasing biological age of both normal and cancer cells[16].

Despite a lot of consistency with the literature, some of our results were less expected. For instance, proteasome inhibitors such as MG-132 and bortezomib have been shown to transiently enhance pluripotency markers early after treatment[100] but to ultimately suppress them at later stages[101]. Given that we identified them as rejuvenating compounds argues that Pasta might particularly capture the earlier effects of perturbations, consistent with CMAP's short exposure duration (~24 hours). Furthermore, these proteasome inhibitors as well as several other of our top rejuvenating hits, including piperlongumine and the structurally related CT-200783, as well as cardiac glycosides such as ouabain—consistent with recent reports[23,102]— and HDAC inhibitors, have previously been described as senolytic agents[68]. Here one may speculate that these compounds do not intrinsically rejuvenate cells but rather eliminate aged or damaged cells from the population.

Among the chemical perturbations identified, pralatrexate emerged as a previously unrecognized senescence-inducing compound. This dihydrofolate inhibitor, approved by the FDA in 2009 for relapsed or refractory peripheral T-cell lymphoma[103], ranked fourth for senescence induction across cell lines and was validated experimentally. Its senescence-inducing effect mirrors those of other dihydrofolate inhibitors such as methotrexate[104–106] (ranked 98th), pemetrexed[107], and pyrimethamine[108]. Notably, pralatrexate ranked first for preferentially inducing senescence in cancer cells, suggesting potential therapeutic selectivity. These results highlight the ability of transcriptomic aging clocks such as Pasta to reveal latent pharmacological properties of established drugs.

We observed heterogeneity in cell line responses to age-modulatory perturbations, such as HDAC inhibitors having pro-aging or pro-rejuvenating effects depending on the cell lines. To better understand the origins of this heterogeneity, we integrated CMAP and DepMap data to identify molecular determinants of the propensity of cells to respond to age-modulatory cues. Transcriptomic and proteomic analyses showed that pro-senescence responses depend strongly on mitochondrial translation, consistent with links between mitochondrial metabolism, reactive oxygen species, and senescence[109]. Mitochondrial translation promotes biogenesis and reactive oxygen species production[110], driving p21-mediated senescence[111], while p53 enhances this process to sustain cytokine output through the SASP[88]. Increased type I interferon gene copy numbers also correlated with stronger aging responses, in line with the pro-aging role of interferon signaling[92,112]. These findings suggest that mitochondrial translation and interferon pathways sensitize cells to senescence via oxidative and inflammatory mechanisms.

Rejuvenation was instead shaped by mRNA splicing and epigenetic factors. Reduced PRC2 and HDAC gene copy numbers heightened sensitivity to rejuvenating stimuli, consistent with PRC2's role as a universal biomarker of aging and rejuvenation[94] and with HDAC inhibition facilitating reprogramming[113]. Together, these results delineate key molecular pathways

underlying propensity for cellular aging and rejuvenation and highlight potential targets for therapeutic modulation.

To conclude, Pasta is a biologically grounded and versatile transcriptomic aging clock with broad utility, spanning fundamental research on human and mouse aging, senescence, and stem cells to clinical diagnostics and therapeutic discovery. By enabling systematic screening of perturbation datasets, it extends the use of conventional aging clocks toward data-driven inference of determinants of cellular biological age. Given the rapid expansion of perturbation datasets, this framework provides a scalable platform for translational research in cancer[16], neurodegeneration[17], regeneration[18–20], and age-related interventions (Extended Data Fig. 10).

Methods

Construction of the aging transcriptomic database

GTEx Protected Access Data version 8 (17,382 samples) was used to obtain precise age information[37]. We discarded 1,674 samples with a RIN score below 6. An additional 1,137 samples originating from gender-specific tissues (cervix, fallopian, uterus, prostate, ovary, testis, vagina), tissues with low sample sizes (e.g., bladder, 12 samples), or non-relevant sources (e.g., *cells_ebv_transformed_lymphocytes*) were further discarded. Raw or normalized count matrices were retrieved in R using the package GEOquery[114] for GEO[22] datasets and using the package ExpressionAtlas[38] for Expression Atlas[115] datasets. The biomaRt R package[116] was used to convert gene or probe IDs into Ensembl gene IDs. Custom scripts were used to parse health status, age, and tissue metadata. Overall, 21 datasets with data from healthy donors were identified. These datasets contained 17,212 samples in total, including 14,571 from GTEx, and 8,113 genes in common, which were also included in the landmark or best inferred gene sets in the CMAP L1000 dataset[13] and were used for further analyses. A rank transformation was applied to each sample using base R's *rank* function with the parameter *ties.method = 'average'*.

Generating sample pairs

For age-shift models, pairs of samples from the same dataset were generated randomly and then filtered. Sample IDs were shuffled 200 times for the training set and 2000 times for the testing set within each dataset to generate IDs for both the first and the second samples in the pairs. Then, pairs with an age difference below the cutoff, involving non-healthy individuals, or representing duplicate pairs (i.e., sample 1 and sample 2 swapped) were excluded. Datasets with more than 1.5 times as many pairs as samples were randomly downsampled to maintain a ratio of 1.5 pairs per sample. Then, the ranked gene expression values were subtracted between sample 1 and sample 2 and used as input to train the models. The output variable was binary, indicating whether sample 1 was younger or older than sample 2.

Training and evaluating models

Ridge-regularized generalized linear models were built using 10-fold cross-validation to identify the optimal lambda parameters. This was done using the function *cv.glmnet* from the glmnet package[117], and using the following parameters: *s_lambda = 'lambda.min'*, and *type.measure = 'mse'* for regression models or *type.measure = 'deviance'* for classification models. Models were compared using a leave-one-dataset-out (LODO) cross-validation

strategy. The age effects of the classification models (traditional and age-shift classifiers) refer to the log-odds obtained using the *predict* function from the glmnet package with parameter *type = 'link'*. The Pearson correlation coefficients of the predicted ages (for the regression model) or age effects (for the classification models) with the true ages in the left-out datasets were used as the metric to compare models. To convert the classifiers' predictions into age differences, referred to as age scores, they were multiplied by a scaling factor specific to each classifier. These scaling factors were obtained by applying the cross-validation models to a set of pairs with balanced age differences in the left-out datasets. Specifically, 1000 pairs were generated per dataset which were then filtered to retain 20 pairs in each of 13 bins representing 10 year age differences. Age predictions were made and results across all datasets were aggregated. Finally, the scaling factor was defined as the beta coefficient from a linear regression of age_difference on age_effect.

Comparing Pasta with MultiTIMER

The MultiTIMER model cannot be used directly, as it requires query samples to be normalized alongside the training samples[30]. As a result, it only functions with samples available on ARCHS4[9]. We re-ran the MultiTIMER training code using ARCHS4 version 2.1, incorporating 14 test datasets —8 from the LODO analysis and 6 novel datasets—alongside the MultiTIMER's training samples to ensure joint batch correction. LODO datasets were evaluated using Pasta's LODO performance, while the novel datasets were assessed using predictions from the final Pasta model.

Extending Pasta for application to mouse samples

Orthologous gene pairs between human and mouse were obtained from the Human Gene Database HGD[118]. The table was filtered to retain only one-to-one orthologues. For each mouse dataset, gene symbols or probes were converted to Ensembl gene identifies. Genes were then restricted to those with one-to-one human orthologues and renamed using human gene names. This yielded 3292 genes, of which 1,600 overlapped with Pasta's predictive genes. Missing genes required for the Pasta age model were imputed with the dataset median, and rank-transformations were conducted before applying the age model.

Conducting gene set enrichment analysis

GSEA was performed using the GSEA function from the *clusterProfiler* R package[119]. Gene sets were retrieved via the msigdbr function from the *msigdbr* package[120], aggregating sets from the CGP, KEGG, REACTOME, and MIRDB subcategories. Model coefficients, sorted by decreasing magnitude, were used as input. The *TERM2GENE* argument was set to the aggregated gene sets, with all other parameters left at default.

Analyzing senescence and stem cell datasets

Supplementary Table 1 from the SenCID manuscript[121] was used to identify samples from senescent, proliferative, and quiescent cells. Custom scripts were used to identify GSM IDs for each sample. This resulted in 308 of 603 samples which could be downloaded and analyzed. For stem cell datasets, samples were manually annotated. Data were downloaded mainly from GEO using Pasta's functions, *getting_geo_count_mat_for_age_prediction* and

getting_GEO_ES_for_age_model, as well as with custom scripts. The Pasta R package was used to make age predictions. AUC-ROC values were computed using the *auc* function from the pROC R package[122].

Analyzing TCGA data

Histological tumor grades were obtained from Ping et al[123] for samples from BRCA patients and from Liu et al[124] for samples from patients with other cancer types, as described in Zheng et al[57]. Only cancer types with histological tumor grades ranging from G1 to G4 were retained, excluding BLCA samples. The TCGAbiolinks R package was used to download gene expression data and obtain age and overall survival information[125]. Only tumor samples were used for all analyses. The Pearson product-moment correlation coefficient test, implemented in base R's cor.test function, was used to compute p-values and confidence intervals for the correlation between Pasta scores or age in days with tumor grades (from 1 to 4). P-values were adjusted for multiple testing using the False Discovery Rate method implemented via base R's *p.adjust* function with *method = 'fdr'*. Survival analyses were conducted by fitting Cox proportional hazards regression models using the *coxph* function from the R survival package[126]. The predictor variable was either standardized Pasta age scores or standardized age in days. Hazard ratios and their confidence intervals were obtained by exponentiating the model's coefficients and their confidence intervals. Kaplan-Meier survival analyses were performed using the *survfit* function from the R survival package, with Pasta's tertiles as the predictor variable.

Analyzing CMAP L1000 data – Preprocessing data and predicting age scores

Raw level 3 data and metadata files were obtained from the https://clue.io website, and MD5 checksums were verified. Several CMAP plates had incorrect or missing annotations, which were corrected using custom scripts. Plates with missing treatment information or fewer than three control samples were excluded from all analyses. The perturbation metadata table was reduced from 3,027,596 to 3,017,946 entries after filtering. Missing mechanisms of action (MOAs) were manually annotated for 60 compounds. Missing compound names were manually annotated for seven compounds that had only BRD (Broad) IDs available. All data were split into individual *.rds* files by plate to facilitate parallel computation with reduced memory usage. For each sample, its final age score was defined as its raw age score minus the mean age score of control samples on the same plate. Control samples were defined as DMSO-treated for compound perturbations and untreated for genetic perturbations.

Analyzing CMAP L1000 data – Annotating senescence- and reprogramming-related compounds

Senescence-inducing compounds were obtained from Table 1 in Petrova et al[61]. Reprogramming-promoting compounds were obtained from Supplementary Table 1 in Knyazer et al[62]. Due to the large number of compounds listed, only those used in at least 3 studies were selected. The selected compounds were manually checked for their presence in the CMAP dataset. In total, 31 senescence-inducing compounds and 11 reprogramming-promoting compounds were identified. The MOA in CMAP for each of these compounds was identified. MOAs for compounds present in both sets were annotated as 'both'. This resulted in 5 MOAs for the 'reprogramming' group, 11 MOAs for the 'senescence' group, and 2

MOAs for the 'both' group. All CMAP compounds belonging to these MOAs were annotated accordingly.

Analyzing CMAP L1000 data – Statistical analysis

Two-sided t-tests were used to assess whether age scores significantly differed from zero for each perturbation, either across all cell lines or within each individual cell line. All dosages, incubation times, and BRD IDs for a given perturbation were pooled to increase statistical power. Bonferroni correction was applied to account for multiple testing. For GPs, a two-stage approach was employed to mitigate noise associated with variability in genetic construct quality. In the first stage, significance was tested for each BRD ID. In the second stage, significance was tested for each GP using only its top two BRD IDs with the highest absolute average score_pv values. Two metrics were created to combine statistical significance with effect size: *score_pv = -log10(p-value) * mean(age score)*, and *score_pa = -log10(adjusted p-value) * mean(age score)*. Significant results were defined as entries with a score_pa greater than or equal to -log10(0.05). Over-representation of MOAs among the significant entries, relative to other entries, was assessed using Fisher's exact test, implemented in base R's fisher.test function with *alternative = 'greater'*. FDR correction was applied to account for multiple testing. Analysis of cancer-specific pro-aging compounds was conducted using one-sided two-sample t-tests to assess whether the mean age in tumor cell lines was greater than that in normal (as defined in the CMAP cell annotation file) cell lines. Only compounds with at least two normal and two tumor cell lines and a mean tumor age effect greater than two were selected. The score_pa and score_pv metrics were derived as described above but using the difference in mean age effect between tumor and normal cell lines as the measure of effect size.

Analyzing CMAP L1000 data – Selecting perturbations and cell lines to display in the heatmaps

Perturbation and cell line selection was performed separately for the Aging and Rejuvenating groups. The perturbation-by-cell line results tables were sorted by decreasing score_pa for the Rejuvenating group and by increasing score_pa for the Aging group. The top 40 perturbations were selected for each group, with a maximum of one perturbation per cell line, which ensured that we displayed the most significant results from our study. For each selected perturbation, the top 5 cell lines were selected.

Analyzing CMAP L1000 data – Pathway enrichment analysis

Significance results from overexpressed genes were inverted to generate a pooled set of GPs with consistent directional effects per gene. The Aging Hallmarks gene sets were obtained from Open Gene[79]. Terms were standardized to the noun forms as these were used for creating the original gene sets. REACTOME gene sets were obtained from the Human MSigDB Collections[46], using the msigdbr R package[120]. Only REACTOME pathways with 300 genes or fewer were retained (1,565 of 1,615). Over-representation analysis was performed using the *enricher* function from the clusterProfiler R package[119] with the parameter *universe* set to all the unique genes in the input database.

Analyzing CMAP L1000 data – Analysis of anticancer drugs

Anticancer drugs were obtained from the CancerDrugs_DB[127] (build date: 28/11/24). We selected drugs approved by the FDA, the EMA, or European national agencies, which retained 311 of 321 drugs. We then kept 132 drugs that were present in the CMAP L1000 dataset. For each cancer, over-representation of its anticancer drugs among the compounds from a given age set (Aging, Rejuvenating or Other), relative to other compounds, was assessed using Fisher's exact test, implemented in base R's fisher.test function with *alternative = 'greater'*. FDR correction was applied to account for multiple testing.

Analyzing CMAP L1000 data – Identifying consistent compound target / gene perturbation results

To identify significant GPs whose age effects are consistent with the targets of significant compounds, we computed the mean score_pa for each gene, using the pooled GP significance table, and for each compound target. Compounds without annotated targets were excluded from this analysis. For compounds with multiple annotated targets, each target was analyzed separately. This resulted in mean score_pa values for 7,638 genes and for 868 targets of compounds. The intersection of these two sets yielded 813 entries, of which 591 had a consistent directionality of effect and were retained. 47 of these had non-zero score_pa values in both sets and were retained. Finally, only perturbation pairs with a min score_pa below 0.05 were retained to include only genes significant in both genetic and chemical perturbations. This resulted in 1 gene for the rejuvenation set (MAP2K1) and 14 genes for the Aging set.

Analyzing DepMap data – Data sources

DepMap data were obtained using the depmap R package[128]. Cell line metadata was retrieved using the *depmap_metadata* function. Drug, CRISPR and RNAi sensitivity data were obtained using the *drug_sensitivity_21Q2*, *depmap_crispr*, and *depmap_rnai* functions, respectively. Copy number, gene expression, protein abundance, and mutation data were obtained using the *depmap_copyNumber*, *depmap_TPM*, *depmap_proteomic*, and *depmap_mutationCalls* functions, respectively.

Analyzing DepMap data – Dependency scores

The significance of age scores for compounds was computed for each BRD ID and cell line, using the same approach as described above (see Statistical analysis). CMAP age score significance results and DepMap dependency scores were then merged by BRD IDs for compounds (41,199 overlapping entries), and by gene names for CRISPR (46,752 overlapping entries) and RNAi (27,401 overlapping entries). The associations between dependency scores and perturbations (compound, CRISPR or RNAi) were assessed by fitting linear models with dependency score as the dependent variable and mean age score as the independent variable. The association between dependency scores and gene sets (MOAs for compounds and REACTOME pathways for CRISPR) was tested by fitting linear models with dependency score as the dependent variable and both mean age score and perturbation name as the independent variables. P-values were adjusted for multiple testing using the Bonferroni correction. This analysis resulted in the identification of 2 MOAs, 2 compounds, 9 REACTOME pathways and no gene perturbations whose mean age scores were significantly associated with dependency scores at an adjusted p-value significance cutoff of 0.05. The

relatively low number of replicates per gene perturbation (median 9, 3rd quartile 12) could explain the lack of significant entries for this analysis.

Analyzing DepMap data – Computing cell lines' propensity for Aging and Rejuvenation

The compound-by-cell line and GP-by-cell line age score significance tables were merged and filtered to retain only perturbations with significant age scores across all cell lines. To derive robust conclusions, the table was further filtered to retain only entries from 19 cell lines in which at least 25 significant perturbations had been tested for both the Rejuvenation and the Aging groups. Propensity scores were then computed as the mean age score of all significant perturbations in the Rejuvenation and Aging groups.

Analyzing DepMap data – Correlating propensity scores with molecular features

The significance of the association between gene mutation and propensity scores was assessed using the Pearson product-moment correlation coefficient test, implemented in base R's *cor.test* function. P-values were corrected for multiple testing using the Bonferroni method. Of 1,724 genes, only TP53 showed a significant association. PCCs between cell lines' propensity for Aging or Rejuvenation and protein expression levels, log-transformed gene copy numbers, or gene expression (in transcripts per million) were computed for each molecule (e.g., gene or protein). Molecules with an absolute PCC higher than 0.4 were analyzed using ORA with the REACTOME database, as described above (see Pathway enrichment analysis).

Cell Culture

A-375 (human melanoma) cells were kindly provided by Dr. Aishe Sarshad (University of Gothenburg, Sweden) and Dr. Alexander Espinosa (Karolinska Institute, Sweden). MDA-MB-231 (human breast cancer) cells were provided by Pr. Kirsty Spalding (Karolinska Institute, Sweden). PC3 cells were kindly provided by Dr. Yvonne Ceder (Lund University). Jurkat cells were purchased from ATCC. Cell lines were maintained in standard DMEM (Gibco). All media were supplemented with 10% heat-inactivated fetal bovine serum (FBS; Gibco) and 1% antibiotics (penicillin/streptomycin 100 U/mL; Gibco). Cells were maintained in a humidified incubator at 37°C and 5% CO2. Cells were tested monthly for *Mycoplasma* contamination using MycoAlert™ Mycoplasma detection kit (Lonza™ LT07-4118-11630271), and only negative cells were used for experiments. For the assessment of pro-aging or rejuvenating effects, cells were incubated in the presence of Pralatrexate (MedChemExpress, #HY-10446, 100 nM in DMSO) or Piperlongumine (MedChemExpress, #HY-N2329, 10 μM) for 72 hours. The DNA damage-inducing agent doxorubicin (D1515, Sigma-Aldrich, 100 nM) was used as a positive control for senescence induction.

SAβG Assay

Cells were washed with PBS, fixed with 0.2% glutaraldehyde for 10 minutes, washed with PBS, and incubated overnight at 37°C with a staining solution containing 1 mg/mL X-Gal (Calbiochem, #203782-1GM) prepared in dimethylformamide (DMF; Sigma-Aldrich, D4551) at pH 6. Cells were then washed in PBS and visualized using an Olympus IX73 brightfield microscope.

Cell Proliferation analysis

Cell proliferation was assessed by EdU incorporation by Click-iT™ Plus EDU Alexa Fluor™ 647 assay (C10634, Thermo Fisher Scientific) as described by the manufacturer. EdU positivity was assessed by Flow Cytometry using a BD FACSCanto II (BD) at the Biomedicum Flow Core Facility. FACS data were analyzed using the FlowJo software.

Immunoblotting

For immunoblotting, proteins extracted by cellular lysis in RIPA buffer were run on 4–12% Bis-Tris acrylamide gels (#NP0322, Thermo Fisher Scientific) and electrotransferred to 0.2 µm polyvinylidene fluoride (PVDF) (#11704156, Biorad) using Trans-Blot Turbo Transfer System (BioRad). Non-specific binding sites were saturated by incubating membranes for 1 h in 0.05% Tween 20 (#P9416, Sigma Aldrich) v-v in Tris-buffered saline (TBS) supplemented with 5% BSA (#10735078001, Sigma-Aldrich, w:v in TBS), followed by an overnight incubation with primary antibodies specific for Waf1/Cip1/CDKN1A p21 (#sc-6246, Santa Cruz Biotechnology), LaminB1 (12987-1-AP, Proteintech) and Oct3/4 (sc-5279, Santa Cruz Biotechnology). Equal protein loading was monitored with GAPDH (#2118, Cell Signaling) and alpha-tubulin (#2144, Cell Signaling) specific antibodies. Membranes were cut in order to allow simultaneous detection of different molecular weight proteins. Membranes were developed with suitable horseradish peroxidase conjugates followed by chemiluminescence-based detection with the Cytiva Amersham ECL Prime (#RPN2232, GE Healthcare) and the Amersham ImageQuant500 software-assisted imager (Cytiva).

Gene Expression Analysis by RT-qPCR

Total RNA was extracted from cell samples using TRIzol reagent (Invitrogen, #15596018) according to the manufacturer's instructions. Up to 1 µg of total RNA was reverse transcribed into cDNA using the iScript cDNA Synthesis Kit (Bio-Rad, #1725038). Quantitative real-time PCR was performed using GoTaq PCR Master Mix (Promega, #A6002) in a QuantStudio 5 Real-Time PCR System (ThermoFisher). The average expression of ACTB (for detection of senescence genes) or GAPDH (for detection of stemness genes) served as endogenous normalization controls. Primers used in this study are the following:

Gene	Forward	Reverse
CDKN1A	CCGCCCCCTCCTCTAGCTGT	CCCCCATCATATACCCCTAACACA
IL6	TACCCCCAGGAGAAGATTCC	TTTTCTGCCAGTGCCTCTTT
OCT4	CCTGAAGCAGAAGAGGATCACC	AAAGCGGCAGATGGTCGTTTGG
SOX2	GCTACAGCATGATGCAGGACCA	TCTGCGAGCTGGTCATGGAGTT
NANOG	TTTGGAAGCTGCTGGGGAAG	GATGGGAGGAGGGGAGAGGA
GAPDH	AATGAAGGGGTCATTGATGG	AAGGTGAAGGTCGGAGTCAA
ACTB	GTTGTCGACGACGAGCG	GCACAGAGCCTCGCCTT

Data availability

All datasets used to train and evaluate the Pasta model are publicly available from the GTEx Portal (https://gtexportal.org), the Gene Expression Omnibus (GEO; https://www.ncbi.nlm.nih.gov/geo/), and the Expression Atlas

(https://www.ebi.ac.uk/gxa/home). Detailed accession numbers and metadata for each dataset are listed in Supplementary Supplementary Table 1. The Cancer Cell Line Encyclopedia (CCLE) and DepMap datasets were accessed via the depmap R package (https://depmap.org/portal/download/). The CMAP L1000 perturbation dataset was obtained from the Connectivity Map website (https://clue.io). The open source Pasta R package, including pre-trained models, functions for data download, processing, and age prediction, is publicly available on GitHub at https://github.com/jsalignon/pasta.

Funding

This work was supported by grants from the Swedish Research Council (VR; 2017-06088, 2019-04868, and 2023-04383), the Swedish Cancer Society (Cancerfonden; 20 1034 Pj and 23 2994 Pj), the Novo Nordisk Foundation (NNF21OC0070427 and NNF22OC0078353), Karolinska Institute (KID2016-00207 and KID2021-00495), by an ICMC project grant and by a Longevity Impetus grant (First round) to C.G.R., and by grants from the Swedish Research Council (MH 2023-02156), the Swedish Cancer Society (24 3543 Pj; 23 0690 JIA), Longevity Impetus Grant from Norn Group (First round), Rosenkranz Foundation and Hevolution Foundation (Third Round), Wellcome Leap's Dynamic Resilience Program (jointly funded by Temasek Trust), Ming Wai Lau Centre for Reparative Medicine (MWLC), Karolinska Institute Senior Researcher Grant to Federico Pietrocola and from Wenner-Gren Stiftelserna (UPD2024-0048) to Federico Pietrocola and Patricia Marques Gonzalez.

Acknowledgements

We thank all members of the Riedel and Pietrocola laboratories for their valuable feedback. We acknowledge the GTEx Consortium, GEO, Expression Atlas, TCGA, CMAP, and DepMap teams for making their transcriptomic and functional genomics datasets publicly available, enabling this research. We thank Sara Hägg (Karolinska Institute) for providing early advice on validation of the transcriptomic aging clock. We thank Bader Zarrouki (AstraZeneca) for scientific discussions.

Conflicts of interest

The authors declare no competing interests.

Author contributions

J.S. performed all computational analyses, developed the R package, generated the computational figures and tables in the manuscript, and wrote the first draft of the manuscript. M.T., P.M.G., and E.R.D. conducted wet lab experiments. F.P., M.T., and P.M.G. generated wet lab figures and associated methods. J.S., H.A., F.P., and C.G.R. contributed to subsequent editing and revision of the manuscript. J.S., F.P., and C.G.R. conceived and supervised the study.

References

1. López-Otín, C., and Kroemer, G. (2021). Hallmarks of Health. Cell *184*, 33–63. https://doi.org/10.1016/j.cell.2020.11.034.

2. López-Otín, C., Blasco, M.A., Partridge, L., Serrano, M., and Kroemer, G. (2023). Hallmarks of aging: An expanding universe. Cell *186*, 243–278. https://doi.org/10.1016/j.cell.2022.11.001.

3. Rosen, R.S., and Yarmush, M.L. (2023). Current Trends in Anti-Aging Strategies. Annual Review of Biomedical Engineering *25*, 363–385. https://doi.org/10.1146/annurev-bioeng-120122-123054.

4. Min, M., Egli, C., Dulai, A.S., and Sivamani, R.K. (2024). Critical review of aging clocks and factors that may influence the pace of aging. Front. Aging *5*. https://doi.org/10.3389/fragi.2024.1487260.

5. Rutledge, J., Oh, H., and Wyss-Coray, T. (2022). Measuring biological age using omics data. Nat Rev Genet, 1–13. https://doi.org/10.1038/s41576-022-00511-7.

6. Teschendorff, A.E., and Horvath, S. (2025). Epigenetic ageing clocks: statistical methods and emerging computational challenges. Nat Rev Genet, 1–19. https://doi.org/10.1038/s41576-024-00807-w.

7. Bird, A. (2002). DNA methylation patterns and epigenetic memory. Genes Dev. *16*, 6–21. https://doi.org/10.1101/gad.947102.

8. Horvath, S., and Raj, K. (2018). DNA methylation-based biomarkers and the epigenetic clock theory of ageing. Nature Reviews Genetics *19*, 371–384. https://doi.org/10.1038/s41576-018-0004-3.

9. Lachmann, A., Torre, D., Keenan, A.B., Jagodnik, K.M., Lee, H.J., Wang, L., Silverstein, M.C., and Ma'ayan, A. (2018). Massive mining of publicly available RNA-seq data from human and mouse. Nat Commun *9*, 1366. https://doi.org/10.1038/s41467-018-03751-6.

10. CZI Cell Science Program, Abdulla, S., Aevermann, B., Assis, P., Badajoz, S., Bell, S.M., Bezzi, E., Cakir, B., Chaffer, J., Chambers, S., et al. (2025). CZ CELLxGENE Discover: a single-cell data platform for scalable exploration, analysis and modeling of aggregated data. Nucleic Acids Research *53*, D886–D900. https://doi.org/10.1093/nar/gkae1142.

11. Youngblut, N.D., Carpenter, C., Prashar, J., Ricci-Tam, C., Ilango, R., Teyssier, N., Konermann, S., Hsu, P.D., Dobin, A., Burke, D.P., et al. (2025). scBaseCamp: an AI agent-curated, uniformly processed, and continually expanding single cell data repository. Preprint at bioRxiv, https://doi.org/10.1101/2025.02.27.640494 https://doi.org/10.1101/2025.02.27.640494.

12. Xu, Z., Wang, W., Yang, T., Li, L., Ma, X., Chen, J., Wang, J., Huang, Y., Gould, J., Lu, H., et al. (2024). STOmicsDB: a comprehensive database for spatial transcriptomics data sharing, analysis and visualization. Nucleic Acids Research *52*, D1053–D1061. https://doi.org/10.1093/nar/gkad933.

13. Subramanian, A., Narayan, R., Corsello, S.M., Peck, D.D., Natoli, T.E., Lu, X., Gould, J., Davis, J.F., Tubelli, A.A., Asiedu, J.K., et al. (2017). A Next Generation Connectivity Map: L1000 Platform and the First 1,000,000 Profiles. Cell *171*, 1437-1452.e17. https://doi.org/10.1016/j.cell.2017.10.049.

14. Wei, Z., Si, D., Duan, B., Gao, Y., Yu, Q., Guo, L., and Liu, Q. (2024). PerturBase: a comprehensive database for single-cell perturbation data analysis and visualization. Preprint at bioRxiv, https://doi.org/10.1101/2024.02.03.578767 https://doi.org/10.1101/2024.02.03.578767.

15. Zhang, J., Ubas, A.A., Borja, R. de, Svensson, V., Thomas, N., Thakar, N., Lai, I., Winters, A., Khan, U., Jones, M.G., et al. (2025). Tahoe-100M: A Giga-Scale Single-Cell Perturbation Atlas for Context-Dependent Gene Function and Cellular Modeling. Preprint at bioRxiv, https://doi.org/10.1101/2025.02.20.639398 https://doi.org/10.1101/2025.02.20.639398.

16. Demaria, M. (2025). Cancer treatments accelerate ageing. Nat Rev Cancer, 1–2. https://doi.org/10.1038/s41568-025-00801-2.

17. Saurat, N., Minotti, A.P., Rahman, M.T., Sikder, T., Zhang, C., Cornacchia, D., Jungverdorben, J., Ciceri, G., Betel, D., and Studer, L. (2024). Genome-wide CRISPR screen identifies neddylation as a regulator of neuronal aging and AD neurodegeneration. Cell Stem Cell *31*, 1162-1174.e8. https://doi.org/10.1016/j.stem.2024.06.001.

18. Abbott, A. (2024). Stem cells head to the clinic: treatments for cancer, diabetes and Parkinson's disease could soon be here. Nature *637*, 18–20. https://doi.org/10.1038/d41586-024-04160-0.

19. Cao, S., and Chen, J. (2025). A speedy cocktail. Nat Chem Biol, 1–2. https://doi.org/10.1038/s41589-024-01804-0.

20. Paine, P.T., Nguyen, A., and Ocampo, A. (2024). Partial cellular reprogramming: A deep dive into an emerging rejuvenation technology. Aging Cell *23*, e14039. https://doi.org/10.1111/acel.14039.

21. Zhang, C., Saurat, N., Cornacchia, D., Chung, S.Y., Sikder, T., Nemchik, A., Minotti, A., Studer, L., and Betel, D. Identifying Age-Modulating Compounds Using a Novel Computational Framework for Evaluating Transcriptional Age. Aging Cell *n/a*, e70075. https://doi.org/10.1111/acel.70075.

22. Barrett, T., Wilhite, S.E., Ledoux, P., Evangelista, C., Kim, I.F., Tomashevsky, M., Marshall, K.A., Phillippy, K.H., Sherman, P.M., Holko, M., et al. (2013). NCBI GEO: archive for functional genomics data sets—update. Nucleic Acids Research *41*, D991–D995. https://doi.org/10.1093/nar/gks1193.

23. Ying, K., Tyshkovskiy, A., Moldakozhayev, A., Wang, H., Magalhaes, C.G.D., Iqbal, S., Garza, A.E., Tskhay, A., Poganik, J.R., Huang, K., et al. (2025). Autonomous AI Agents Discover Aging Interventions from Millions of Molecular Profiles. Preprint at bioRxiv, https://doi.org/10.1101/2023.02.28.530532 https://doi.org/10.1101/2023.02.28.530532.

24. Peters, M.J., Joehanes, R., Pilling, L.C., Schurmann, C., Conneely, K.N., Powell, J., Reinmaa, E., Sutphin, G.L., Zhernakova, A., Schramm, K., et al. (2015). The transcriptional landscape of age in human peripheral blood. Nature Communications *6*, 8570. https://doi.org/10.1038/ncomms9570.

25. Mamoshina, P., Volosnikova, M., Ozerov, I.V., Putin, E., Skibina, E., Cortese, F., and Zhavoronkov, A. (2018). Machine Learning on Human Muscle Transcriptomic Data for Biomarker Discovery and Tissue-Specific Drug Target Identification. Frontiers in Genetics *9*.

26. Fleischer, J.G., Schulte, R., Tsai, H.H., Tyagi, S., Ibarra, A., Shokhirev, M.N., Huang, L., Hetzer, M.W., and Navlakha, S. (2018). Predicting age from the transcriptome of human dermal fibroblasts. Genome Biology *19*, 221. https://doi.org/10.1186/s13059-018-1599-6.

27. Holzscheck, N., Falckenhayn, C., Söhle, J., Kristof, B., Siegner, R., Werner, A., Schössow, J., Jürgens, C., Völzke, H., Wenck, H., et al. (2021). Modeling transcriptomic age using knowledge-primed artificial neural networks. npj Aging Mech Dis *7*, 1–13. https://doi.org/10.1038/s41514-021-00068-5.

28. Shokhirev, M.N., and Johnson, A.A. (2021). Modeling the human aging transcriptome across tissues, health status, and sex. Aging Cell *20*, e13280. https://doi.org/10.1111/acel.13280.

29. Ren, X., and Kuan, P.F. (2020). RNAAgeCalc: A multi-tissue transcriptional age calculator. PLOS ONE *15*, e0237006. https://doi.org/10.1371/journal.pone.0237006.

30. Jung, S., Arcos Hodar, J., and del Sol, A. (2023). Measuring biological age using a functionally interpretable multi-tissue RNA clock. Aging Cell *22*, e13799. https://doi.org/10.1111/acel.13799.

31. Bulteau, R., and Francesconi, M. (2022). Real age prediction from the transcriptome with RAPToR. Nat Methods *19*, 969–975. https://doi.org/10.1038/s41592-022-01540-0.

32. Qi, H., Zhao, H., Li, E., Lu, X., Yu, N., Liu, J., and Han, J. (2025). DeepQA: A Unified Transcriptome-Based Aging Clock Using Deep Neural Networks. Aging Cell *24*, e14471. https://doi.org/10.1111/acel.14471.

33. Zakar-Polyák, E., Csordas, A., Pálovics, R., and Kerepesi, C. (2024). Profiling the transcriptomic age of single-cells in humans. Commun Biol *7*, 1–15. https://doi.org/10.1038/s42003-024-07094-5.

34. Schmidt, Y.G., Wu, D., Madan, S., Sinha, S., Sahni, S., Gopalan, V., Wang, B., Dhruba, S.R., Schäffer, A.A., Weng, N., et al. (2024). IMMClock reveals immune aging and T cell function at single-cell resolution. Preprint at bioRxiv, https://doi.org/10.1101/2024.11.13.623449 https://doi.org/10.1101/2024.11.13.623449.

35. Li, W., Zhang, Z., Kumar, S., Botey-Bataller, J., Zoodsma, M., Ehsani, A., Zhan, Q., Alaswad, A., Zhou, L., Grondman, I., et al. (2025). Single-cell immune aging clocks reveal inter-individual heterogeneity during infection and vaccination. Nat Aging, 1–15. https://doi.org/10.1038/s43587-025-00819-z.

36. Zhu, H., Chen, J., Liu, K., Gao, L., Wu, H., Ma, L., Zhou, J., Liu, Z., and Han, J.-D.J. (2023). Human PBMC scRNA-seq–based aging clocks reveal ribosome to inflammation balance as a single-cell aging hallmark and super longevity. Science Advances *9*, eabq7599. https://doi.org/10.1126/sciadv.abq7599.

37. THE GTEX CONSORTIUM (2020). The GTEx Consortium atlas of genetic regulatory effects across human tissues. Science *369*, 1318–1330. https://doi.org/10.1126/science.aaz1776.

38. Keays, M. (2023). ExpressionAtlas: Download datasets from EMBL-EBI Expression Atlas https://doi.org/10.18129/B9.bioc.ExpressionAtlas.

39. Wetzel, S.J., Ryczko, K., Melko, R.G., and Tamblyn, I. (2022). Twin Neural Network Regression. Applied AI Letters *3*, e78. https://doi.org/10.1002/ail2.78.

40. Wetzel, S.J., Melko, R.G., and Tamblyn, I. (2022). Twin neural network regression is a semi-supervised regression algorithm. Mach. Learn.: Sci. Technol. *3*, 045007. https://doi.org/10.1088/2632-2153/ac9885.

41. Tynes, M., Gao, W., Burrill, D.J., Batista, E.R., Perez, D., Yang, P., and Lubbers, N. (2021). Pairwise Difference Regression: A Machine Learning Meta-algorithm for Improved Prediction and Uncertainty Quantification in Chemical Search. J. Chem. Inf. Model. *61*, 3846–3857. https://doi.org/10.1021/acs.jcim.1c00670.

42. Belaid, M.K., Rabus, M., and Hüllermeier, E. (2024). Pairwise Difference Learning for Classification. Preprint at arXiv, https://doi.org/10.48550/arXiv.2406.20031 https://doi.org/10.48550/arXiv.2406.20031.

43. Wang, A., Chiou, J., Poirion, O.B., Buchanan, J., Valdez, M.J., Verheyden, J.M., Hou, X., Kudtarkar, P., Narendra, S., Newsome, J.M., et al. (2020). Single-cell multiomic profiling of human lungs reveals cell-type-specific and age-dynamic control of SARS-CoV2 host genes. eLife *9*, e62522. https://doi.org/10.7554/eLife.62522.

44. Lukowski, S.W., Lo, C.Y., Sharov, A.A., Nguyen, Q., Fang, L., Hung, S.S., Zhu, L., Zhang, T., Grünert, U., Nguyen, T., et al. (2019). A single□cell transcriptome atlas of the adult human retina. The EMBO Journal *38*, e100811. https://doi.org/10.15252/embj.2018100811.

45. Solé-Boldo, L., Raddatz, G., Schütz, S., Mallm, J.-P., Rippe, K., Lonsdorf, A.S., Rodríguez-Paredes, M., and Lyko, F. (2020). Single-cell transcriptomes of the human skin reveal age-related loss of fibroblast priming. Commun Biol *3*, 1–12. https://doi.org/10.1038/s42003-020-0922-4.

46. Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S., et al. (2005). Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. Proceedings of the National Academy of Sciences *102*, 15545–15550. https://doi.org/10.1073/pnas.0506580102.

47. Stewart-Ornstein, J., Iwamoto, Y., Miller, M.A., Prytyskach, M.A., Ferretti, S., Holzer, P., Kallen, J., Furet, P., Jambhekar, A., Forrester, W.C., et al. (2021). p53

dynamics vary between tissues and are linked with radiation sensitivity. Nat Commun *12*, 898. https://doi.org/10.1038/s41467-021-21145-z.

48. Park, S.-Y., Lee, J.H., Ha, M., Nam, J.-W., and Kim, V.N. (2009). miR-29 miRNAs activate p53 by targeting p85α and CDC42. Nat Struct Mol Biol *16*, 23–29. https://doi.org/10.1038/nsmb.1533.

49. Swahari, V., Nakamura, A., Hollville, E., Hung, Y.-H., Kanke, M., Kurtz, C.L., Caravia, X.M., Roiz-Valle, D., He, S., Krishnamurthy, J., et al. (2024). miR-29 is an important driver of aging-related phenotypes. Commun Biol *7*, 1–11. https://doi.org/10.1038/s42003-024-06735-z.

50. Mandic, R., Marquardt, A., Terhorst, P., Ali, U., Nowak-Rossmann, A., Cai, C., Rodepeter, F.R., Stiewe, T., Wezorke, B., Wanzel, M., et al. (2022). The importin beta superfamily member RanBP17 exhibits a role in cell proliferation and is associated with improved survival of patients with HPV+ HNSCC. BMC Cancer *22*, 785. https://doi.org/10.1186/s12885-022-09854-0.

51. Mertens, J., Paquola, A.C.M., Ku, M., Hatch, E., Böhnke, L., Ladjevardi, S., McGrath, S., Campbell, B., Lee, H., Herdy, J.R., et al. (2015). Directly Reprogrammed Human Neurons Retain Aging-Associated Transcriptomic Signatures and Reveal Age-Related Nucleocytoplasmic Defects. Cell Stem Cell *17*, 705–718. https://doi.org/10.1016/j.stem.2015.09.001.

52. Ye, W., Ling, S., Liu, R.-Y., Pan, Z.-Z., Wang, G., Gao, S., Wu, J., Cao, L., Dong, L., Li, Y., et al. (2018). Exome sequencing reveals the genetic landscape and frequent inactivation of PCDHB3 in Chinese rectal cancers. The Journal of Pathology *245*, 222–234. https://doi.org/10.1002/path.5073.

53. Boutelle, A.M., Mabene, A.R., Yao, D., Xu, H., Wang, M., Tang, Y.J., Lopez, S.S., Sinha, S., Demeter, J., Cheng, R., et al. (2025). Integrative multiomic approaches reveal ZMAT3 and p21 as conserved hubs in the p53 tumor suppression network. Cell Death Differ *32*, 1954–1969. https://doi.org/10.1038/s41418-025-01513-8.

54. Srivatsan, S.R., McFaline-Figueroa, J.L., Ramani, V., Saunders, L., Cao, J., Packer, J., Pliner, H.A., Jackson, D.L., Daza, R.M., Christiansen, L., et al. (2020). Massively multiplex chemical transcriptomics at single-cell resolution. Science *367*, 45–51. https://doi.org/10.1126/science.aax6234.

55. Wang, L., Leite de Oliveira, R., Wang, C., Fernandes Neto, J.M., Mainardi, S., Evers, B., Lieftink, C., Morris, B., Jochems, F., Willemsen, L., et al. (2017). High-Throughput Functional Genetic and Compound Screens Identify Targets for Senescence Induction in Cancer. Cell Reports *21*, 773–783. https://doi.org/10.1016/j.celrep.2017.09.085.

56. Malta, T.M., Sokolov, A., Gentles, A.J., Burzykowski, T., Poisson, L., Weinstein, J.N., Kamińska, B., Huelsken, J., Omberg, L., Gevaert, O., et al. (2018). Machine Learning Identifies Stemness Features Associated with Oncogenic Dedifferentiation. Cell *173*, 338-354.e15. https://doi.org/10.1016/j.cell.2018.03.034.

57. Zheng, H., Song, K., Fu, Y., You, T., Yang, J., Guo, W., Wang, K., Jin, L., Gu, Y., Qi, L., et al. (2021). An absolute human stemness index associated with oncogenic

dedifferentiation. Briefings in Bioinformatics *22*, 2151–2160. https://doi.org/10.1093/bib/bbz174.

58. Schmitt, C.A., Wang, B., and Demaria, M. (2022). Senescence and cancer - role and therapeutic opportunities. Nat Rev Clin Oncol *19*, 619–636. https://doi.org/10.1038/s41571-022-00668-4.

59. López-Otín, C., Pietrocola, F., Roiz-Valle, D., Galluzzi, L., and Kroemer, G. (2023). Meta-hallmarks of aging and cancer. Cell Metab *35*, 12–35. https://doi.org/10.1016/j.cmet.2022.11.001.

60. Colucci, M., Sarill, M., Maddalena, M., Valdata, A., Troiani, M., Massarotti, M., Bolis, M., Bressan, S., Kohl, A., Robesti, D., et al. (2025). Senescence in cancer. Cancer Cell *43*, 1204–1226. https://doi.org/10.1016/j.ccell.2025.05.015.

61. Petrova, N.V., Velichko, A.K., Razin, S.V., and Kantidze, O.L. (2016). Small molecule compounds that induce cellular senescence. Aging Cell *15*, 999–1017. https://doi.org/10.1111/acel.12518.

62. Knyazer, A., Bunu, G., Toren, D., Mracica, T.B., Segev, Y., Wolfson, M., Muradian, K.K., Tacutu, R., and Fraifeld, V.E. (2021). Small molecules for cell reprogramming: a systems biology analysis. Aging (Albany NY) *13*, 25739–25762. https://doi.org/10.18632/aging.203791.

63. Huangfu, D., Maehr, R., Guo, W., Eijkelenboom, A., Snitow, M., Chen, A.E., and Melton, D.A. (2008). Induction of pluripotent stem cells by defined factors is greatly improved by small-molecule compounds. Nat Biotechnol *26*, 795–797. https://doi.org/10.1038/nbt1418.

64. Anokye-Danso, F., Trivedi, C.M., Juhr, D., Gupta, M., Cui, Z., Tian, Y., Zhang, Y., Yang, W., Gruber, P.J., Epstein, J.A., et al. (2011). Highly Efficient miRNA-Mediated Reprogramming of Mouse and Human Somatic Cells to Pluripotency. Cell Stem Cell *8*, 376–388. https://doi.org/10.1016/j.stem.2011.03.001.

65. Di Bernardo, G., Squillaro, T., Dell'Aversana, C., Miceli, M., Cipollaro, M., Cascino, A., Altucci, L., and Galderisi, U. (2009). Histone Deacetylase Inhibitors Promote Apoptosis and Senescence in Human Mesenchymal Stem Cells. Stem Cells and Development *18*, 573–582. https://doi.org/10.1089/scd.2008.0172.

66. Li, X.-N., Shu, Q., Su, J.M.-F., Perlaky, L., Blaney, S.M., and Lau, C.C. (2005). Valproic acid induces growth arrest, apoptosis, and senescence in medulloblastomas by increasing histone hyperacetylation and regulating expression of p21Cip1, CDK4, and CMYC. Molecular Cancer Therapeutics *4*, 1912–1922. https://doi.org/10.1158/1535-7163.MCT-05-0184.

67. An, H.-M., Xue, Y.-F., Shen, Y.-L., Du, Q., and Hu, B. (2013). Sodium Valproate Induces Cell Senescence in Human Hepatocarcinoma Cells. Molecules *18*, 14935–14947. https://doi.org/10.3390/molecules181214935.

68. Rad, A.N., and Grillari, J. (2024). Current senolytics: Mode of action, efficacy and limitations, and their future. Mechanisms of Ageing and Development *217*, 111888. https://doi.org/10.1016/j.mad.2023.111888.

69. Wang, Y., Wang, X., Liu, L., and Wang, X. (2009). HDAC inhibitor trichostatin A-inhibited survival of dopaminergic neuronal cells. Neuroscience Letters *467*, 212–216. https://doi.org/10.1016/j.neulet.2009.10.037.

70. Zhu, H., Blake, S., Kusuma, F.K., Pearson, R.B., Kang, J., and Chan, K.T. (2020). Oncogene-induced senescence: From biology to therapy. Mechanisms of Ageing and Development *187*, 111229. https://doi.org/10.1016/j.mad.2020.111229.

71. Xu, S., Wu, W., Huang, H., Huang, R., Xie, L., Su, A., Liu, S., Zheng, R., Yuan, Y., Zheng, H., et al. (2019). The p53/miRNAs/Ccna2 pathway serves as a novel regulator of cellular senescence: Complement of the canonical p53/p21 pathway. Aging Cell *18*, e12918. https://doi.org/10.1111/acel.12918.

72. Fang, L., Li, D., Yin, J., Pan, H., Ye, H., Bowman, J., Capaldo, B., and Kelly, K. (2022). TMPRSS2-ERG promotes the initiation of prostate cancer by suppressing oncogene-induced senescence. Cancer Gene Ther *29*, 1463–1476. https://doi.org/10.1038/s41417-022-00454-5.

73. Zhuang, D., Mannava, S., Grachtchouk, V., Tang, W.-H., Patil, S., Wawrzyniak, J.A., Berman, A.E., Giordano, T.J., Prochownik, E.V., Soengas, M.S., et al. (2008). C-MYC overexpression is required for continuous suppression of oncogene-induced senescence in melanoma cells. Oncogene *27*, 6623–6634. https://doi.org/10.1038/onc.2008.258.

74. Guney, I., Wu, S., and Sedivy, J.M. (2006). Reduced c-Myc signaling triggers telomere-independent senescence by regulating Bmi-1 and p16INK4a. Proceedings of the National Academy of Sciences *103*, 3645–3650. https://doi.org/10.1073/pnas.0600069103.

75. Takahashi, K., and Yamanaka, S. (2006). Induction of Pluripotent Stem Cells from Mouse Embryonic and Adult Fibroblast Cultures by Defined Factors. Cell *126*, 663–676. https://doi.org/10.1016/j.cell.2006.07.024.

76. Balcerek, J., Jiang, J., Li, Y., Jiang, Q., Holdreith, N., Singh, B., Chandra, V., Lv, K., Ren, J., Rozenova, K., et al. (2018). Lnk/Sh2b3 deficiency restores hematopoietic stem cell function and genome integrity in Fancd2 deficient Fanconi anemia. Nat Commun *9*, 3915. https://doi.org/10.1038/s41467-018-06380-1.

77. Song, H., Wang, J., Wang, X., Yuan, B., Li, D., Hu, A., Guo, Y., Cai, S., Jin, S., Zhou, Y., et al. (2022). HNF4A-AS1-encoded small peptide promotes self-renewal and aggressiveness of neuroblastoma stem cells via eEF1A1-repressed SMAD4 transactivation. Oncogene *41*, 2505–2519. https://doi.org/10.1038/s41388-022-02271-4.

78. Baugh, E.H., Ke, H., Levine, A.J., Bonneau, R.A., and Chan, C.S. (2018). Why are there hotspot mutations in the TP53 gene in human cancers? Cell Death Differ *25*, 154–160. https://doi.org/10.1038/cdd.2017.180.

79. Rafikova, E., Nemirovich-Danchenko, N., Ogmen, A., Parfenenkova, A., Velikanova, A., Tikhonov, S., Peshkin, L., Rafikov, K., Spiridonova, O., Belova, Y., et al. (2024). Open

Genes—a new comprehensive database of human genes associated with aging and longevity. Nucleic Acids Research *52*, D950–D962. https://doi.org/10.1093/nar/gkad712.

80. Milacic, M., Beavers, D., Conley, P., Gong, C., Gillespie, M., Griss, J., Haw, R., Jassal, B., Matthews, L., May, B., et al. (2024). The Reactome Pathway Knowledgebase 2024. Nucleic Acids Research *52*, D672–D678. https://doi.org/10.1093/nar/gkad1025.

81. Sagara, A., Miura, S., Kobinata, A., Naganawa, R., Yaginuma, S., Saito, S., Saito, R., Kominato, H., Yumoto, T., and Sato, F. (2023). COL8A1 enhances the invasion/metastasis in MDA-MB-231 cells via the induction of *IL1B* and *MMP1* expression. Biochemical and Biophysical Research Communications *642*, 145–153. https://doi.org/10.1016/j.bbrc.2022.12.046.

82. Sugrue, M.M., Shin, D.Y., Lee, S.W., and Aaronson, S.A. (1997). Wild-type p53 triggers a rapid senescence program in human tumor cells lacking functional p53. Proceedings of the National Academy of Sciences *94*, 9648–9653. https://doi.org/10.1073/pnas.94.18.9648.

83. Rufini, A., Tucci, P., Celardo, I., and Melino, G. (2013). Senescence and aging: the critical roles of p53. Oncogene *32*, 5129–5143. https://doi.org/10.1038/onc.2012.640.

84. Jiang, Z., Zhu, M., Zhang, L., Cui, H., Jiang, R., and Yang, Y. (2024). Antitumor immunity and prognosis value elicited by FAT3 and LRP1B co-mutation in endometrial cancer. Gynecologic Oncology *187*, 1–11. https://doi.org/10.1016/j.ygyno.2024.04.023.

85. Schneider, T., Martinez-Martinez, A., Cubillos-Rojas, M., Bartrons, R., Ventura, F., and Rosa, J.L. (2019). Large HERCs Function as Tumor Suppressors. Front. Oncol. *9*. https://doi.org/10.3389/fonc.2019.00524.

86. Namasu, C.Y., Katzerke, C., Bräuer-Hartmann, D., Wurm, A.A., Gerloff, D., Hartmann, J.-U., Schwind, S., Müller-Tidow, C., Hilger, N., Fricke, S., et al. (2017). ABR, a novel inducer of transcription factor C/EBPα, contributes to myeloid differentiation and is a favorable prognostic factor in acute myeloid leukemia. Oncotarget *8*, 103626–103639. https://doi.org/10.18632/oncotarget.22093.

87. Full article: Role of Rho guanine nucleotide exchange factors in non-small cell lung cancer https://www.tandfonline.com/doi/full/10.1080/21655979.2021.2006519?rfr_dat=cr_pub++0pubmed&url_ver=Z39.88-2003&rfr_id=ori%3Arid%3Acrossref.org.

88. Jiang, X., Baig, A.H., Palazzo, G., Del Pizzo, R., Bortecen, T., Groessl, S., Zaal, E.A., Amaya Ramirez, C.C., Kowar, A., Aviles-Huerta, D., et al. (2024). P53-dependent hypusination of eIF5A affects mitochondrial translation and senescence immune surveillance. Nat Commun *15*, 7458. https://doi.org/10.1038/s41467-024-51901-w.

89. Wang, B., Han, J., Elisseeff, J.H., and Demaria, M. (2024). The senescence-associated secretory phenotype and its physiological and pathological implications. Nat Rev Mol Cell Biol *25*, 958–978. https://doi.org/10.1038/s41580-024-00727-x.

90. Yu, Q., Katlinskaya, Y.V., Carbone, C.J., Zhao, B., Katlinski, K.V., Zheng, H., Guha, M., Li, N., Chen, Q., Yang, T., et al. (2015). DNA-Damage-Induced Type I Interferon

Promotes Senescence and Inhibits Stem Cell Function. Cell Reports *11*, 785–797. https://doi.org/10.1016/j.celrep.2015.03.069.

91. Katlinskaya, Y.V., Katlinski, K.V., Yu, Q., Ortiz, A., Beiting, D.P., Brice, A., Davar, D., Sanders, C., Kirkwood, J.M., Rui, H., et al. (2016). Suppression of Type I Interferon Signaling Overcomes Oncogene-Induced Senescence and Mediates Melanoma Development and Progression. Cell Reports *15*, 171–180. https://doi.org/10.1016/j.celrep.2016.03.006.

92. Sun, E.D., Zhou, O.Y., Hauptschein, M., Rappoport, N., Xu, L., Navarro Negredo, P., Liu, L., Rando, T.A., Zou, J., and Brunet, A. (2024). Spatial transcriptomic clocks reveal cell proximity effects in brain ageing. Nature, 1–12. https://doi.org/10.1038/s41586-024-08334-8.

93. Pereira, C.F., Piccolo, F.M., Tsubouchi, T., Sauer, S., Ryan, N.K., Bruno, L., Landeira, D., Santos, J., Banito, A., Gil, J., et al. (2010). ESCs Require PRC2 to Direct the Successful Reprogramming of Differentiated Cells toward Pluripotency. Cell Stem Cell *6*, 547–556. https://doi.org/10.1016/j.stem.2010.04.013.

94. Moqri, M., Cipriano, A., Simpson, D.J., Rasouli, S., Murty, T., de Jong, T.A., Nachun, D., de Sena Brandine, G., Ying, K., Tarkhov, A., et al. (2024). PRC2-AgeIndex as a universal biomarker of aging and rejuvenation. Nat Commun *15*, 5956. https://doi.org/10.1038/s41467-024-50098-2.

95. Koldobskiy, M.A., Camacho, O., Reddy, P., Belmonte, J.C.I., and Feinberg, A.P. (2023). Convergence of aging- and rejuvenation-related epigenetic alterations on PRC2 targets. Preprint at bioRxiv, https://doi.org/10.1101/2023.06.08.544045 https://doi.org/10.1101/2023.06.08.544045.

96. Salignon, J., Millan-Ariño, L., Garcia, M.U., and Riedel, C.G. (2024). Cactus: A user-friendly and reproducible ATAC-Seq and mRNA-Seq analysis pipeline for data preprocessing, differential analysis, and enrichment analysis. Genomics *116*, 110858. https://doi.org/10.1016/j.ygeno.2024.110858.

97. Li, H., Collado, M., Villasante, A., Strati, K., Ortega, S., Cañamero, M., Blasco, M.A., and Serrano, M. (2009). The Ink4/Arf locus is a barrier for iPS cell reprogramming. Nature *460*, 1136–1139. https://doi.org/10.1038/nature08290.

98. Rufini, A., Tucci, P., Celardo, I., and Melino, G. (2013). Senescence and aging: the critical roles of p53. Oncogene *32*, 5129–5143. https://doi.org/10.1038/onc.2012.640.

99. Wang, L., Lankhorst, L., and Bernards, R. (2022). Exploiting senescence for the treatment of cancer. Nat Rev Cancer *22*, 340–355. https://doi.org/10.1038/s41568-022-00450-9.

100. Floyd, E.Z., Staszkiewicz, J., Power, R.A., Kilroy, G., Kirk-Ballard, H., Barnes, C.W., Strickler, K.L., Rim, J.S., Harkins, L.L., Gao, R., et al. (2015). Prolonged Proteasome Inhibition Cyclically Upregulates Oct3/4 and Nanog Gene Expression, but Reduces Induced Pluripotent Stem Cell Colony Formation. Cell Reprogram *17*, 95–105. https://doi.org/10.1089/cell.2014.0030.

101. Selenina, A.V., Tsimokha, A.S., and Tomilin, A.N. (2017). Proteasomes in Protein Homeostasis of Pluripotent Stem Cells. Acta Naturae *9*, 39–47.

102. Tyshkovskiy, A., Kholdina, D., Ying, K., Davitadze, M., Molière, A., Tongu, Y., Kasahara, T., Kats, L.M., Vladimirova, A., Moldakozhayev, A., et al. (2024). Transcriptomic Hallmarks of Mortality Reveal Universal and Specific Mechanisms of Aging, Chronic Disease, and Rejuvenation. Preprint at bioRxiv, https://doi.org/10.1101/2024.07.04.601982 https://doi.org/10.1101/2024.07.04.601982.

103. O'Connor, O.A., Amengual ,Jennifer, Colbourn ,Donald, Deng ,Changchun, and and Sawas, A. (2017). Pralatrexate: a comprehensive update on pharmacology, clinical activity and strategies to optimize use. Leukemia & Lymphoma *58*, 2548–2557. https://doi.org/10.1080/10428194.2017.1306642.

104. Hattangadi, D.K., DeMasters, G.A., Walker, T.D., Jones, K.R., Di, X., Newsham, I.F., and Gewirtz, D.A. (2004). Influence of p53 and caspase 3 activity on cell death and senescence in response to methotrexate in the breast tumor cell. Biochemical Pharmacology *68*, 1699–1708. https://doi.org/10.1016/j.bcp.2004.06.033.

105. Dabrowska, M., Mosieniak, G., Skierski, J., Sikora, E., and Rode, W. (2009). Methotrexate-induced senescence in human adenocarcinoma cells is accompanied by induction of p21waf1/cip1 expression and lack of polyploidy. Cancer Letters *284*, 95–101. https://doi.org/10.1016/j.canlet.2009.04.015.

106. Csik, B., Vali Kordestan, K., Gulej, R., Patai, R., Nyul-Toth, A., Shanmugarama, S., Mukli, P., Ungvari, A., Balsara, K.E., McNall, R.Y., et al. (2025). Cisplatin and methotrexate induce brain microvascular endothelial and microglial senescence in mouse models of chemotherapy-associated cognitive impairment. GeroScience. https://doi.org/10.1007/s11357-025-01560-6.

107. Chen, Y., Zhang, C., Jin, S., Li, J., Dai, J., Zhang, Z., and Guo, R. (2023). Pemetrexed induces ROS generation and cellular senescence by attenuating TS-mediated thymidylate metabolism to reverse gefitinib resistance in NSCLC. Journal of Cellular and Molecular Medicine *27*, 2032–2044. https://doi.org/10.1111/jcmm.17799.

108. Dong, H., Hu, L., Li, W., Shi, M., He, L., Wang, C., Hu, Y., Wang, H., Wen, C., Liu, H., et al. (2022). Pyrimethamine inhibits cell growth by inducing cell senescence and boosting CD8+ T-cell mediated cytotoxicity in colorectal cancer. Mol Biol Rep *49*, 4281–4292. https://doi.org/10.1007/s11033-022-07262-y.

109. Sun, N., Youle, R.J., and Finkel, T. (2016). The Mitochondrial Basis of Aging. Molecular Cell *61*, 654–666. https://doi.org/10.1016/j.molcel.2016.01.028.

110. Abu Shelbayeh, O., Arroum, T., Morris, S., and Busch, K.B. (2023). PGC-1α Is a Master Regulator of Mitochondrial Lifecycle and ROS Stress Response. Antioxidants (Basel) *12*, 1075. https://doi.org/10.3390/antiox12051075.

111. Macip, S., Igarashi, M., Fang, L., Chen, A., Pan, Z., Lee, S.W., and Aaronson, S.A. (2002). Inhibition of p21□mediated ROS accumulation can rescue p21□induced senescence. The EMBO Journal *21*, 2180–2188. https://doi.org/10.1093/emboj/21.9.2180.

112. Frisch, S.M., and MacFawn, I.P. (2020). Type I interferons and related pathways in cell senescence. Aging Cell *19*, e13234. https://doi.org/10.1111/acel.13234.

113. Pereira, B., Correia, F.P., Alves, I.A., Costa, M., Gameiro, M., Martins, A.P., and Saraiva, J.A. (2024). Epigenetic reprogramming as a key to reverse ageing and increase longevity. Ageing Research Reviews *95*, 102204. https://doi.org/10.1016/j.arr.2024.102204.

114. Davis, S., and Meltzer, P.S. (2007). GEOquery: a bridge between the Gene Expression Omnibus (GEO) and BioConductor. Bioinformatics *23*, 1846–1847. https://doi.org/10.1093/bioinformatics/btm254.

115. George, N., Fexova, S., Fuentes, A.M., Madrigal, P., Bi, Y., Iqbal, H., Kumbham, U., Nolte, N.F., Zhao, L., Thanki, A.S., et al. (2024). Expression Atlas update: insights from sequencing data at both bulk and single cell level. Nucleic Acids Research *52*, D107–D114. https://doi.org/10.1093/nar/gkad1021.

116. Durinck, S., Spellman, P.T., Birney, E., and Huber, W. (2009). Mapping identifiers for the integration of genomic datasets with the R/Bioconductor package biomaRt. Nature Protocols *4*, 1184–1191.

117. Friedman, J.H., Hastie, T., and Tibshirani, R. (2010). Regularization Paths for Generalized Linear Models via Coordinate Descent. Journal of Statistical Software *33*, 1–22. https://doi.org/10.18637/jss.v033.i01.

118. Duan, G., Wu, G., Chen, X., Tian, D., Li, Z., Sun, Y., Du, Z., Hao, L., Song, S., Gao, Y., et al. (2023). HGD: an integrated homologous gene database across multiple species. Nucleic Acids Research *51*, D994–D1002. https://doi.org/10.1093/nar/gkac970.

119. Wu, T., Hu, E., Xu, S., Chen, M., Guo, P., Dai, Z., Feng, T., Zhou, L., Tang, W., Zhan, L., et al. (2021). clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. The Innovation *2*, 100141. https://doi.org/10.1016/j.xinn.2021.100141.

120. Dolgalev, I. (2022). msigdbr: MSigDB Gene Sets for Multiple Organisms in a Tidy Data Format.

121. Tao, W., Yu, Z., and Han, J.-D.J. (2024). Single-cell senescence identification reveals senescence heterogeneity, trajectory, and modulators. Cell Metabolism *36*, 1126-1143.e5. https://doi.org/10.1016/j.cmet.2024.03.009.

122. Robin, X., Turck, N., Hainard, A., Tiberti, N., Lisacek, F., Sanchez, J.-C., and Müller, M. (2011). pROC: an open-source package for R and S+ to analyze and compare ROC curves. BMC Bioinformatics *12*, 77. https://doi.org/10.1186/1471-2105-12-77.

123. Ping, Z., Xia, Y., Shen, T., Parekh, V., Siegal, G.P., Eltoum, I.-E., He, J., Chen, D., Deng, M., Xi, R., et al. (2016). A microscopic landscape of the invasive breast cancer genome. Sci Rep *6*, 27545. https://doi.org/10.1038/srep27545.

124. Liu, J., Lichtenberg, T., Hoadley, K.A., Poisson, L.M., Lazar, A.J., Cherniack, A.D., Kovatich, A.J., Benz, C.C., Levine, D.A., Lee, A.V., et al. (2018). An Integrated TCGA Pan-Cancer Clinical Data Resource to Drive High-Quality Survival Outcome Analytics. Cell *173*, 400-416.e11. https://doi.org/10.1016/j.cell.2018.02.052.

125. Colaprico, A., Silva, T.C., Olsen, C., Garofano, L., Cava, C., Garolini, D., Sabedot, T., Malta, T.M., Pagnotta, S.M., Castiglioni, I., et al. (2015). TCGAbiolinks: An R/Bioconductor package for integrative analysis of TCGA data. Nucleic Acids Research. https://doi.org/10.1093/nar/gkv1507.

126. Therneau, T.M. (2024). A Package for Survival Analysis in R.

127. Pantziarka, P., Capistrano I, R., De Potter, A., Vandeborne, L., and Bouche, G. (2021). An Open Access Database of Licensed Cancer Drugs. Front Pharmacol *12*, 627574. https://doi.org/10.3389/fphar.2021.627574.

128. Gatto, L., and Killian, T. (2023). depmap: Cancer Dependency Map Data Package https://doi.org/10.18129/B9.bioc.depmap.

129. Kim, Y.-M., Byun, H.-O., Jee, B.A., Cho, H., Seo, Y.-H., Kim, Y.-S., Park, M.H., Chung, H.-Y., Woo, H.G., and Yoon, G. (2013). Implications of time-series gene expression profiles of replicative senescence. Aging Cell *12*, 622–634. https://doi.org/10.1111/acel.12087.

130. Liu, X., Ouyang, J.F., Rossello, F.J., Tan, J.P., Davidson, K.C., Valdes, D.S., Schröder, J., Sun, Y.B.Y., Chen, J., Knaupp, A.S., et al. (2020). Reprogramming roadmap reveals route to human induced trophoblast stem cells. Nature *586*, 101–107. https://doi.org/10.1038/s41586-020-2734-6.

131. Li, Q., Hutchins, A.P., Chen, Y., Li, S., Shan, Y., Liao, B., Zheng, D., Shi, X., Li, Y., Chan, W.-Y., et al. (2017). A sequential EMT-MET mechanism drives the differentiation of human embryonic stem cells towards hepatocytes. Nat Commun *8*, 15166. https://doi.org/10.1038/ncomms15166.

Figure legends

Figure 1. The age-shift classifier enables accurate cross-platform prediction of relative transcriptomic age.

(**a**) Overview of the evaluated modelling strategies. Models were trained and tested exclusively on samples from healthy donors. (**b-d**) Schematic diagrams of the regression (**b**), traditional classifier (**c**), and age-shift classifier (**d**) strategies. (**e**) Comparative performance in predicting relative age, assessed using leave-one-dataset-out cross-validation. All samples were rank transformed except for REG_RC and AS40_RC where raw counts were used. The y-axis shows the Pearson Correlation Coefficient (PCC) between chronological ages and predicted age scores in the held-out datasets. (**f-h**) Age predictions for single-cell RNA-Seq aging atlases of the retina[44] (**f**), skin[45] (**g**), and lung[43] (**h**). Abbreviations: REG, regression; AS30/40/50, age-shift classifier with 30, 40, or 50 year age difference; TC_30_70 and TC_40_60, traditional classifiers with young cutoffs at 30 or 40 years and old cutoffs at 70 or 60 years; RC, raw counts.

Figure 2. Pasta predicts transcriptomic age across tissues through p53-linked signals.

(**a-b**) Pearson Correlation Coefficient (PCC) of age scores for Pasta and MultiTIMER compared with chronological ages for healthy individuals. (**a**) Results from the leave-one-dataset-out analysis. (**b**) Results on external validation datasets. (**c-d**) PCC of age scores for Pasta compared with chronological age for mouse datasets GSE40156 (**c**) and GSE167665 (**d**). (**e**) Gene set enrichment analysis (GSEA) of Pasta coefficients using the REACTOME, CGP, KEGG and MIRDB gene sets from the Human MSigDB Collections. The grey lines mark the adjusted p-value cutoff of 0.05. (**f-g**) Top ten genes with the most positive (**f**) and negative (**g**) coefficients.

Figure 3. Pasta identifies senescent and stem cells and may help stratify tumor grade and patient overall survival.

(**a-b**) Pasta's ability to discriminate between senescent and proliferative (**a**) or quiescent (**b**) cells, assessed using AUC-ROC (Area Under the Curve of the Receiver Operating Characteristic) values, across multiple RNA-Seq datasets. (**c**) Pasta's predictions on a microarray senescence time course dataset[129]. (**d**) Pasta was applied to pseudobulks from a SciPlex dataset generated by exposing the A549 cell line to 138 drugs[54], and the resulting age scores were evaluated for their capacity to recover senescence-inducing drugs, as reported in Wang et al 2017[55]. Only the highest drug concentration (10 µM) was used for this analysis. (**e-f**) Pasta's ability to discriminate between stem cells (**e**) or induced-pluripotent stem cells (iPSCs) (**f**) and fibroblasts, assessed by AUC-ROC. (**g**) Pasta's age scores on a bulk RNA-Seq OSKM-mediated reprogramming time-course dataset[130]. (**h**) Pasta's age scores on a bulk RNA-Seq liver differentiation time-course dataset[131]. (**i**) Correlations of patient chronological ages (red) and Pasta's age scores (blue) with tumor grade. (**j-k**) Boxplots of Pasta's age scores for breast invasive carcinoma (**j**) and kidney renal clear cell carcinoma (**k**) patients, grouped by tumor grade. (**l**) Hazard ratios from Cox proportional-hazards regression models fitted for scaled ages (red) and scaled age scores (blue). (**m-n**) Kaplan-Meier survival estimates for adrenocortical carcinoma (**m**) and thymoma (**n**) patients, grouped by age score tertiles.

Figure 4. Pasta identifies age-modulatory compounds.

(**a-d**) Yellow, blue, green, and black indicate drug classes or compounds from drug classes that induce senescence[61], promote reprogramming[62], affect both processes, or neither, respectively. (**a**) Volcano plot showing the most significant Rejuvenating compounds with negative age scores and Aging compounds with positive age scores. For display purposes, p-values equal to zero were set to 10^{-300}. (**b**) Proportion of compounds annotated as senescence-inducing, pro-reprogramming, both, or neither for significantly age-increasing (Aging), age-decreasing (Rejuvenating), and other (Other) compounds. (**c-d**) Significantly enriched mechanisms of action (MOA) for Aging (**c**) and Rejuvenating (**d**) compounds. The grey lines mark the adjusted p-value cutoff of 0.05. (**e**) Heatmap of compounds by cell line showing the most significant age scores (see Methods). Colors indicate the square root of a metric balancing significance and effect size: $-log_{10}(adjusted\ p\text{-}value) \times mean(age\ score)$. Missing values are shown in grey.

Figure 5. Experimental validation of pralatrexate as a senescence-inducer and piperlongumine as a rejuvenating agent.

(**a-h**) Testing senescence-induction by pralatrexate in responder A375 and non-responder MDA-MB-231 cells. (**a,e**) Representative images from Senescence Associated Beta Galactosidase assays in A375 (**a**) and MDA-MB-231 (**e**) cells exposed to pralatrexate. Doxorubicin served as the positive control for senescence induction. Scale bar 10 micrometers. (**b,f**) qPCR analysis of CDKN1A and IL6 mRNA levels in A375 (**b**) and MDA-MB-231 (**f**) cells treated with pralatrexate or doxorubicin. N=3 biological replicates; *p< 0.05; *** p< 0.001; one-way ANOVA. (**c,g**) Cell proliferation by EdU incorporation in A375 (**c**) and MDA-MB-231 (**g**) cells after treatment with pralatrexate or doxorubicin. One representative experiment with three replicates per condition. ** p< 0.01; *** p< 0.001 (one-way ANOVA, calculated on EdU^+ cells). (**d,h**) Immunoblotting showing LMNB1, CDKN1A and GAPDH protein levels in A375 (**d**) and MDA-MB-231 (**h**) cells after treatment with pralatrexate or doxorubicin. N=3 biological replicates. (**i-l**) Testing stemness-induction of piperlongumine in responder PC3 and non-responder MCF7 cells. (**i,k**) qPCR analysis of OCT4, SOX2 and NANOG levels in PC3 (**i**) and MCF7 (**k**) cells exposed to piperlongumine. * p< 0.05; Student's t test. (**j,l**) Immunoblotting showing OCT3/4 and alpha-tubulin protein levels in PC3 (**j**) and MCF7 (**l**) cells treated with piperlongumine. Abbreviations: Prltx, pralatrexate; Doxo, doxorubicin; Piplng, piperlongumine; ns, non-significant.

Figure 6. Pasta identifies age-modulatory gene perturbations.

(**a**) Volcano plot displaying the most significant Rejuvenating gene perturbations (GPs) with negative age scores and Aging GPs with positive age scores. (**b-f**) Enrichment analysis using the Hallmarks of Aging[79] (**b-c**) and REACTOME (**d-f**) gene sets for Rejuvenating GPs (**b, d**), Aging GPs (**c, e**), and Aging GPs that are also compound targets of significant Aging compounds (**f**). Significant results from the three GP types (OE, KD, and KO) were pooled, with overexpression results inverted (i.e., MYC OE was included in the Aging GP set). Grey lines mark the adjusted p-value cutoff of 0.05. (**g**) Heatmap of GPs by cell lines showing the most significant age scores (see Methods). Colors indicate the square root of a metric balancing significance and effect size: $-log_{10}(adjusted\ p\text{-}value) \times mean(age\ score)$. Missing values are shown in grey.

Figure 7. Pasta uncovers molecular determinants of cellular propensity for aging and rejuvenation.

(**a**) Propensity of various cell lines to shift toward older or younger states following exposure to age-regulatory perturbations (see Methods). (**b**) Gene mutation counts most strongly negatively correlated with the propensity for Aging. (**c**) TP53 mutation count versus propensity for Aging. (**d-i**) REACTOME pathway enrichment for genes whose protein abundance (**d, g, i**), mRNA expression (**e**), and copy numbers (**f, h**) exhibit an absolute PCC higher than 0.4 with the propensity of cells for Rejuvenation (**d, e, f**) and Aging (**g, h, i**). The grey lines mark the adjusted p-value cutoff of 0.05.

Supplementary Figures legends

Extended Data Fig. 1. Results of leave-one-dataset-out (LODO) analysis.

(**a**) Comparative performance in predicting relative age, assessed with LODO, when considering only young (< 40 years old) and old (> 60 years old) individuals. (**b**) Detailed LODO results for the AS40/Pasta model when considering all individuals. The x-axis shows the AUC-ROC for predicting whether the first sample is younger than the second in each evaluation pair.

Extended Data Fig. 2. Predicted versus true age differences for sample pairs using the AS40/Pasta LODO models.

For each dataset, predictions were generated by a model trained on all other datasets. Each point corresponds to a sample pair with an age difference of at least 40 years.

Extended Data Fig. 3. Predicted age scores versus true ages for all samples using the AS40/Pasta LODO models.

For each dataset, predictions were generated by a model trained on all other datasets. Each point corresponds to one sample.

Extended Data Fig. 4. Baseline regression model identifies fewer known age-modulatory compounds in L1000 data.

Proportion of compounds annotated as senescence-inducing, pro-reprogramming, both, or neither for significantly age-increasing (Aging), age-decreasing (Rejuvenating) compounds, and other compounds (Other). Analyses were performed identically to Figure 4b but employed the baseline regression model from Figure 1 rather than the age-shift model.

Extended Data Fig. 5. Chemotherapeutics broadly elevate transcriptomic age.

(**a**) Cancers whose anticancer drugs are significantly enriched in age-modulatory compounds across all cell lines. Only cancers with at least 5 approved anticancer drugs according to CancerDrugs_DB[127] were selected. 17 of 27 cancers were significant for the Aging group. The grey line marks the adjusted p-value cutoff of 0.05. (**b**) Top 10 compounds whose age scores are statistically higher in cancer cell lines than in normal cell lines (see Methods). (**c**) Mean age score in cancer and normal cell lines for the 10 compounds shown in (b).

Extended Data Fig. 6. Relationship between compound-induced changes in cellular age and viability.

(**a-c**) Yellow, blue, and black indicate compounds in the Aging, Rejuvenating, and Other groups, respectively. (**a**) Boxplots and violin plots of dependency scores for compounds by cell line in the Aging (175 entries), Rejuvenating (75 entries), and Other (40,906 entries) groups. (**b-c**) Mechanisms of action (**b**) and compounds (**c**) with the strongest associations between age scores and dependency. (**d**) Mean age score versus dependency for MDM inhibitors. Each point represents a compound in a given cell line.

Extended Data Fig. 7. Relationship between gene perturbation-induced changes in cellular age and viability.

(**a-b**) Yellow, blue, and black indicate gene perturbations (GPs) in the Aging, Rejuvenating, and Other groups, respectively. (**a**) Boxplots and violin plots of dependency scores for GPs by cell line in the Aging, Rejuvenating, and Other groups. Number of entries per set: Other (KD: 27,243; KO: 46,684 entries), Aging (KD: 152; KO: 55), Rejuvenating (KD: 6; KO: 13). (**b**) REACTOME pathways whose PCCs are most significantly associated with dependency scores in the knockout Aging group. No term was enriched in the Rejuvenating group. The grey lines mark the 0.05 adjusted p-value cutoff. (**c-d**) Mean age score versus dependency for GPs in the retinoic acid (RA) biosynthesis (**c**) and in the oncogene-induced senescence (**d**) pathways. Each point represents a gene knockout in a given cell line.

Extended Data Fig. 8. REACTOME enrichment analysis of age-modulatory gene perturbations.

Results for Rejuvenating knockdown (**a**), knockout (**b**), and overexpression (**c**), and for Aging knockdown (**d**), knockout (**e**), and overexpression (**f**). (**g**) Genes with non-zero adjusted p-value scores (score_pa) for their age-effect across cell types that showed consistency between the pooled gene perturbations and compound target analyses.

Extended Data Fig. 9. REACTOME enrichment analysis of pathways significantly associated with cellular propensity for Aging and Rejuvenation.

Results for genes whose copy number positively correlates with the propensity for Rejuvenation (**a**), proteins whose abundance negatively correlates with the propensity for Aging (**b**), and genes whose transcript levels positively correlate with the propensity for Rejuvenation (**c**) or negatively correlate with the propensity for Rejuvenation (**d**).

Extended Data Fig. 10. Graphical summary of the study.

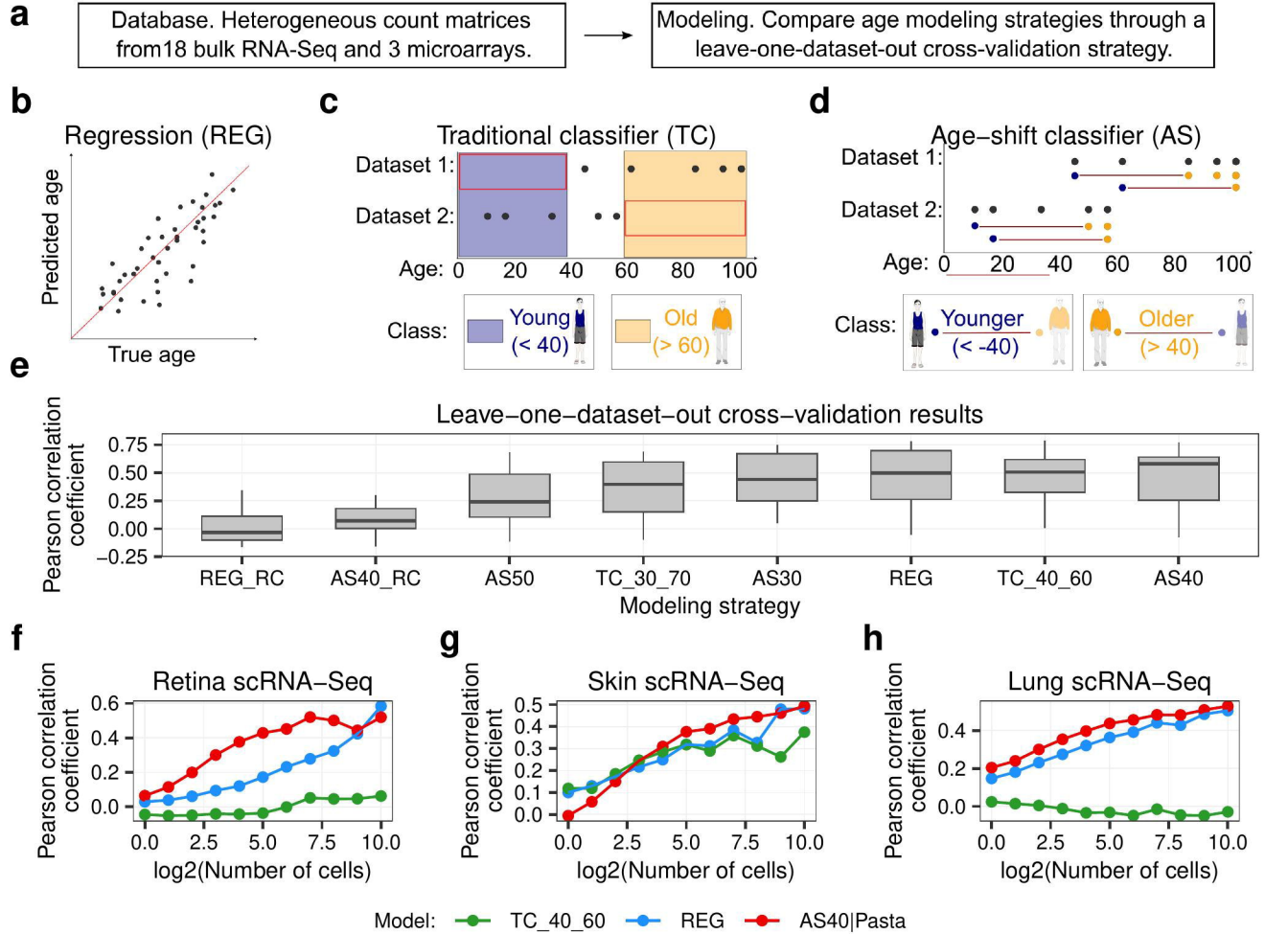

a
Database. Heterogeneous count matrices from18 bulk RNA-Seq and 3 microarrays.
Modeling. Compare age modeling strategies through a leave-one-dataset-out cross-validation strategy.
b
Regression (REG)
Predicted age
True age
c
Traditional classifier (TC)
Dataset 1:
Dataset 2:
Age: 0 20 40 60 80 100
Class:
Young (< 40)
Old (> 60)
d
Age−shift classifier (AS)
Dataset 1:
Dataset 2:
Age: 0 20 40 60 80 100
Class:
Younger (< -40)
Older (> 40)
e
Leave−one−dataset−out cross−validation results
Pearson correlation coefficient
0.75
0.50
0.25
0.00
−0.25
REG_RC
AS40_RC
AS50
TC_30_70
AS30
REG
TC_40_60
AS40
Modeling strategy
f
Retina scRNA−Seq
Pearson correlation coefficient
0.6
0.4
0.2
0.0
0.0 2.5 5.0 7.5 10.0
log2(Number of cells)
g
Skin scRNA−Seq
Pearson correlation coefficient
0.5
0.4
0.3
0.2
0.1
0.0
0.0 2.5 5.0 7.5 10.0
log2(Number of cells)
h
Lung scRNA−Seq
Pearson correlation coefficient
0.4
0.2
0.0
0.0 2.5 5.0 7.5 10.0
log2(Number of cells)
Model: TC_40_60 REG AS40|Pasta

Figure 1

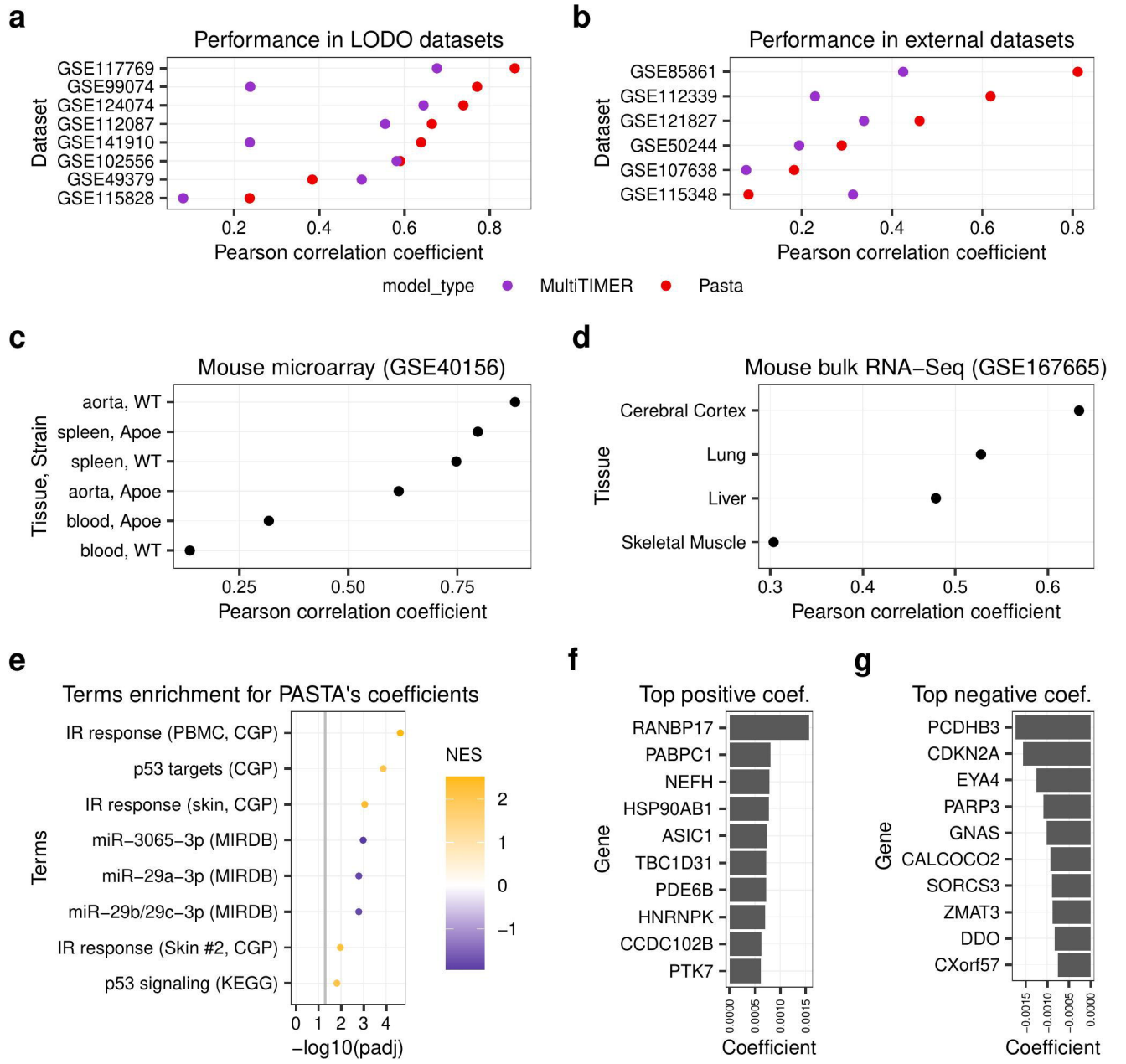

a
Performance in LODO datasets
Dataset
GSE117769
GSE99074
GSE124074
GSE112087
GSE141910
GSE102556
GSE49379
GSE115828
0.2
0.4
0.6
0.8
Pearson correlation coefficient
b
Performance in external datasets
Dataset
GSE85861
GSE112339
GSE121827
GSE50244
GSE107638
GSE115348
0.2
0.4
0.6
0.8
Pearson correlation coefficient
model_type
MultiTIMER
Pasta
c
Mouse microarray (GSE40156)
Tissue, Strain
aorta, WT
spleen, Apoe
spleen, WT
aorta, Apoe
blood, Apoe
blood, WT
0.25
0.50
0.75
Pearson correlation coefficient
d
Mouse bulk RNA−Seq (GSE167665)
Tissue
Cerebral Cortex
Lung
Liver
Skeletal Muscle
0.3
0.4
0.5
0.6
Pearson correlation coefficient
e
Terms enrichment for PASTA's coefficients
Terms
IR response (PBMC, CGP)
p53 targets (CGP)
IR response (skin, CGP)
miR−3065−3p (MIRDB)
miR−29a−3p (MIRDB)
miR−29b/29c−3p (MIRDB)
IR response (Skin #2, CGP)
p53 signaling (KEGG)
0
1
2
3
4
−log10(padj)
NES
2
1
0
−1
f
Top positive coef.
Gene
RANBP17
PABPC1
NEFH
HSP90AB1
ASIC1
TBC1D31
PDE6B
HNRNPK
CCDC102B
PTK7
0.0000
0.0005
0.0010
0.0015
Coefficient
g
Top negative coef.
Gene
PCDHB3
CDKN2A
EYA4
PARP3
GNAS
CALCOCO2
SORCS3
ZMAT3
DDO
CXorf57
−0.0015
−0.0010
−0.0005
0.0000
Coefficient

Figure 2

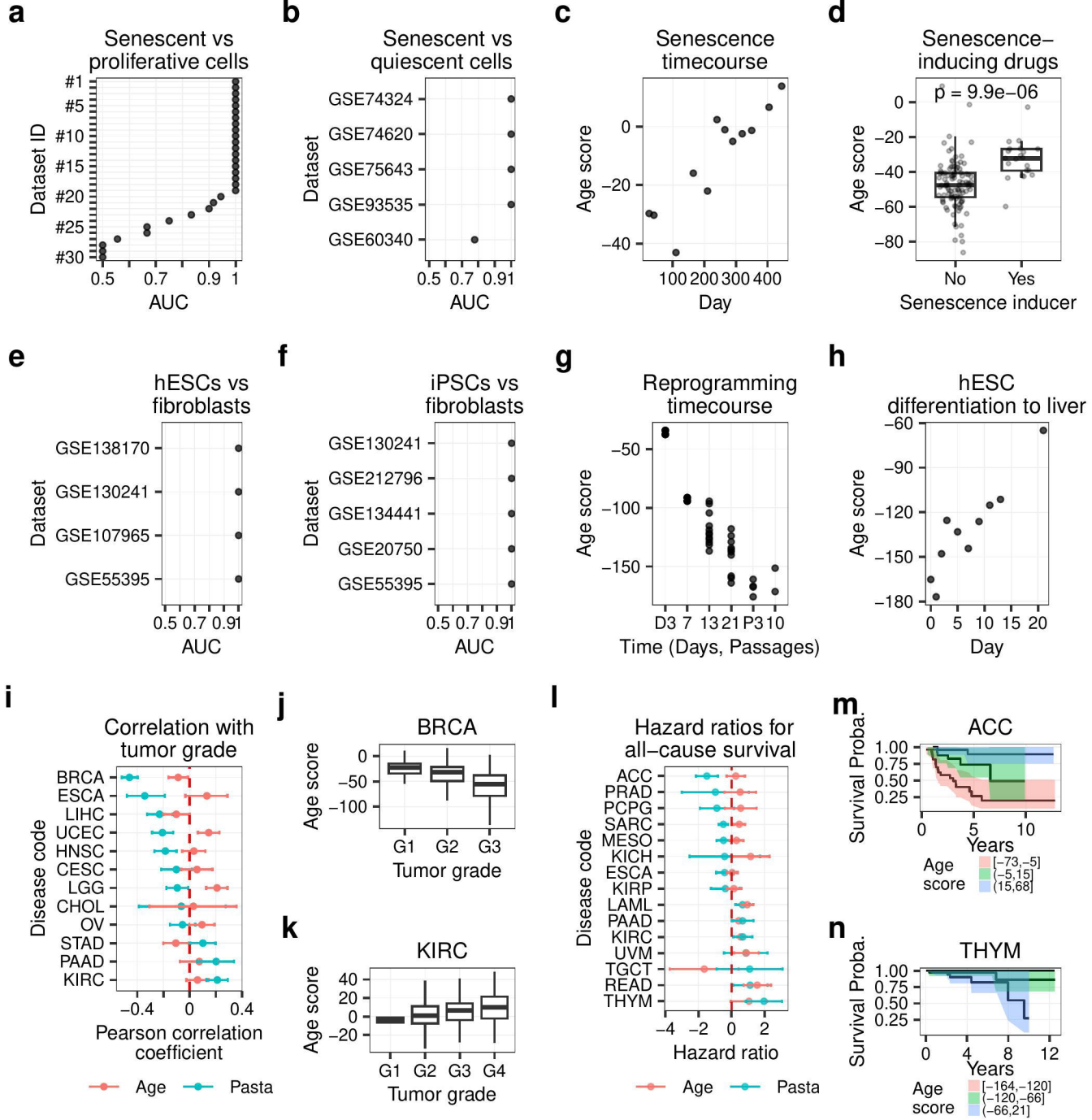

a
Senescent vs proliferative cells
Dataset ID
#1
#5
#10
#15
#20
#25
#30
0.5
0.7
0.9
1
AUC
b
Senescent vs quiescent cells
Dataset
GSE74324
GSE74620
GSE75643
GSE93535
GSE60340
0.5 0.7 0.9 1
AUC
c
Senescence timecourse
Age score
0
−20
−40
100 200 300 400
Day
d
Senescence–inducing drugs
p = 9.9e−06
Age score
0
−20
−40
−60
−80
No
Yes
Senescence inducer
e
hESCs vs fibroblasts
Dataset
GSE138170
GSE130241
GSE107965
GSE55395
0.5 0.7 0.9 1
AUC
f
iPSCs vs fibroblasts
Dataset
GSE130241
GSE212796
GSE134441
GSE20750
GSE55395
0.5 0.7 0.9 1
AUC
g
Reprogramming timecourse
Age score
−50
−100
−150
D3 7 13 21 P3 10
Time (Days, Passages)
h
hESC differentiation to liver
Age score
−60
−90
−120
−150
−180
0 5 10 15 20
Day
i
Correlation with tumor grade
Disease code
BRCA
ESCA
LIHC
UCEC
HNSC
CESC
LGG
CHOL
OV
STAD
PAAD
KIRC
−0.4
0
0.4
Pearson correlation coefficient
Age
Pasta
j
BRCA
Age score
0
−50
−100
G1
G2
G3
Tumor grade
k
KIRC
Age score
40
20
0
−20
G1
G2
G3
G4
Tumor grade
l
Hazard ratios for all–cause survival
Disease code
ACC
PRAD
PCPG
SARC
MESO
KICH
ESCA
KIRP
LAML
PAAD
KIRC
UVM
TGCT
READ
THYM
−4
−2
0
2
Hazard ratio
Age
Pasta
m
ACC
Survival Proba.
1.00
0.75
0.50
0.25
0
5
10
Years
Age score
[−73,−5]
(−5,15]
(15,68]
n
THYM
Survival Proba.
1.00
0.75
0.50
0.25
0
4
8
12
Years
Age score
[−164,−120]
(−120,−66]
(−66,21]

Figure 3

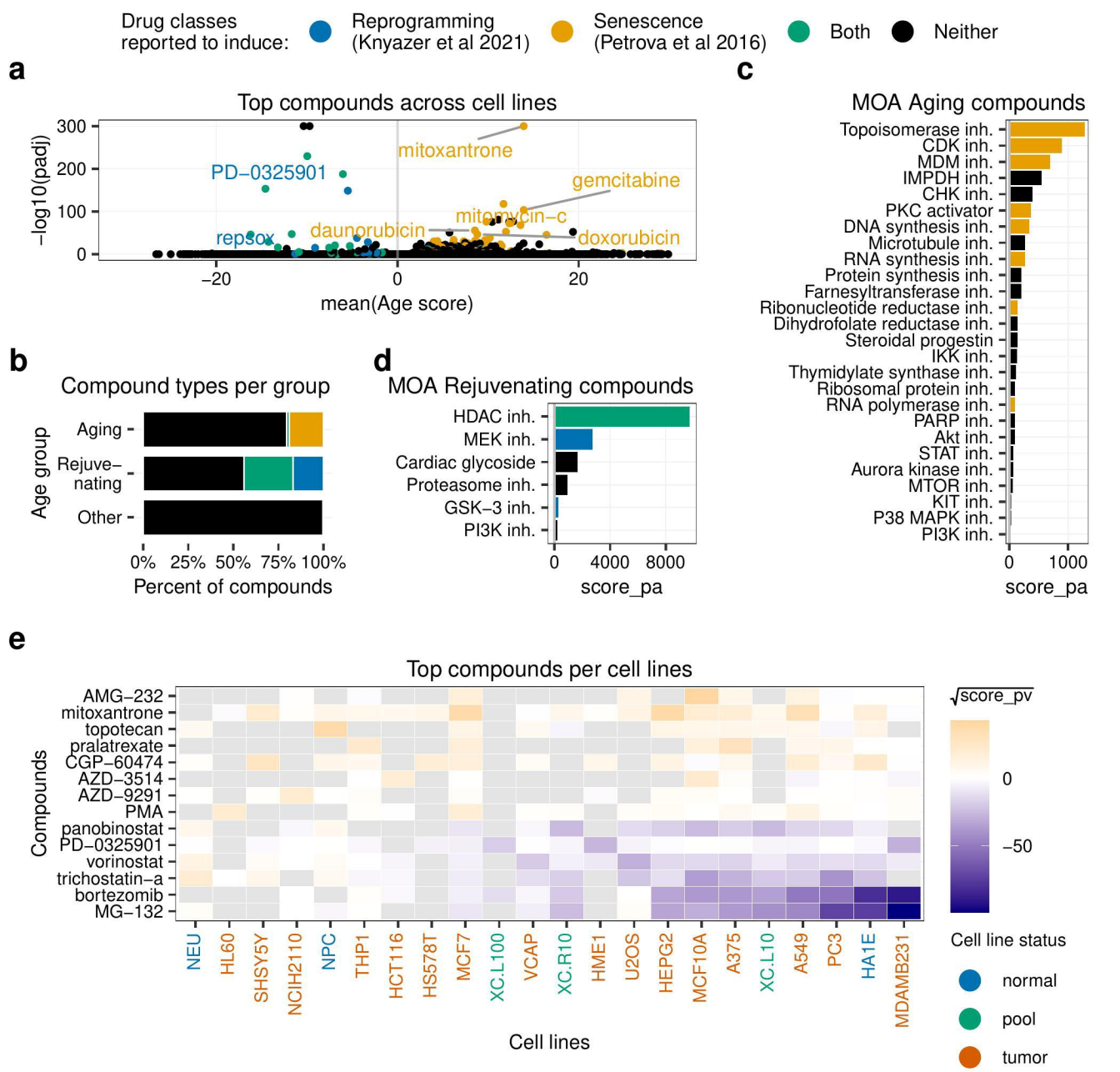
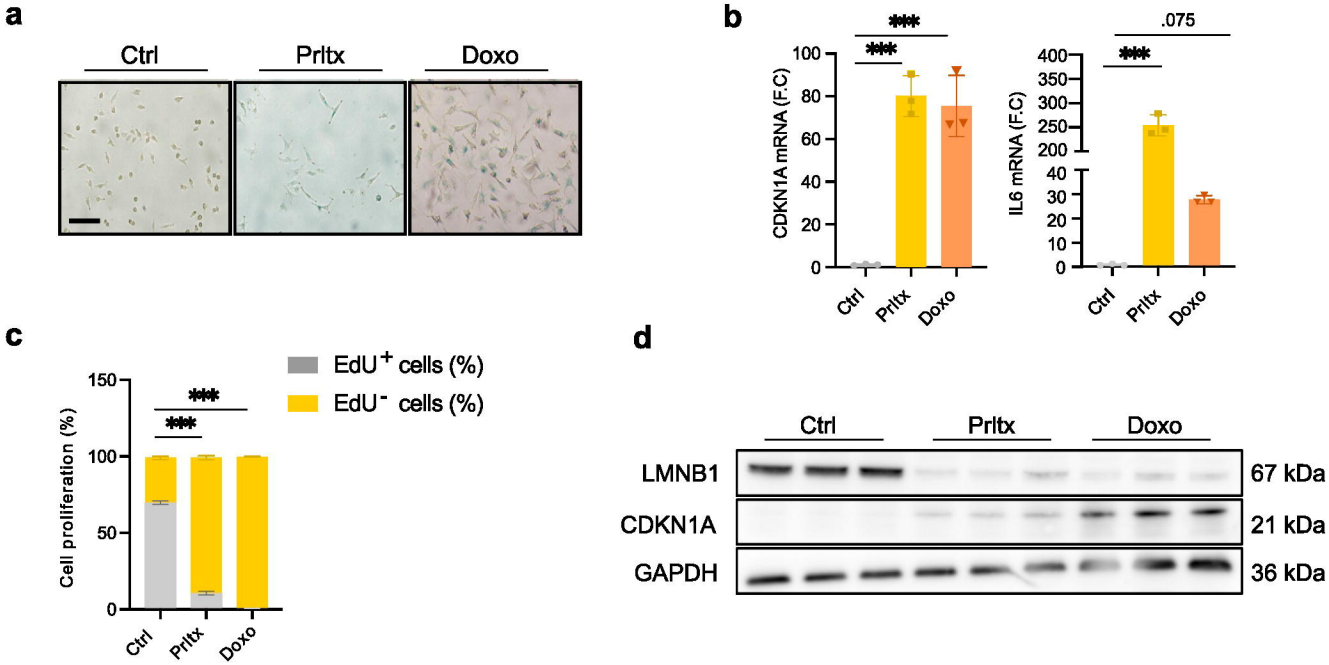
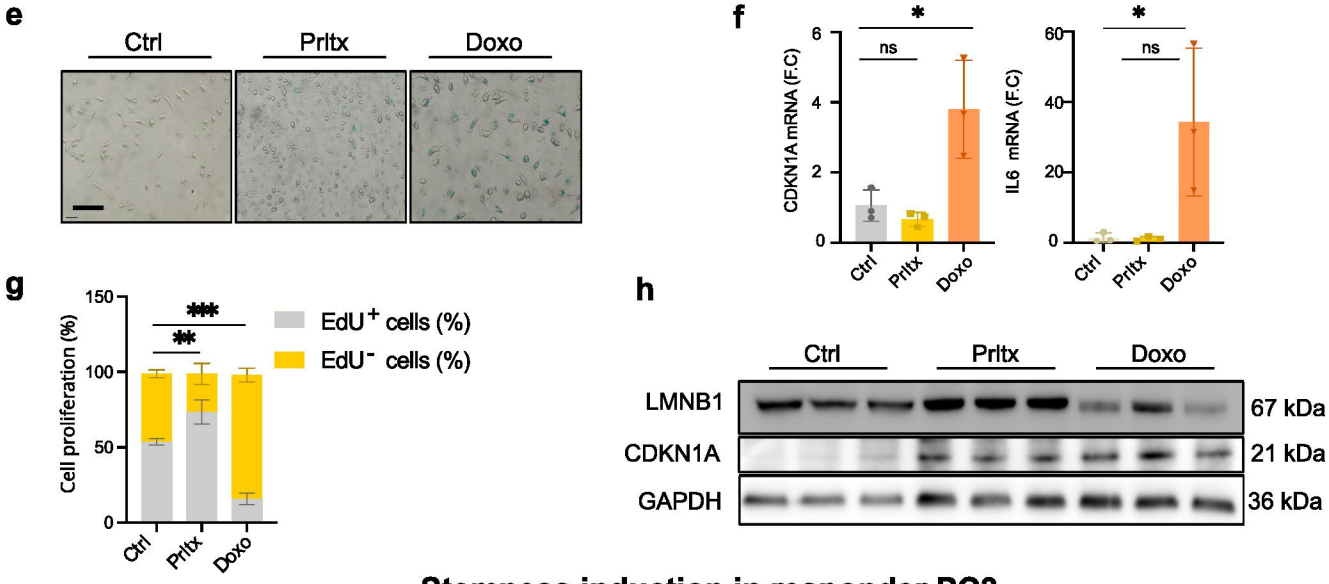

Drug classes reported to induce:
Reprogramming (Knyazer et al 2021)
Senescence (Petrova et al 2016)
Both
Neither
a
Top compounds across cell lines
−log10(padj)
mean(Age score)
mitoxantrone
PD−0325901
gemcitabine
mitomycin−c
repsox
daunorubicin
doxorubicin
b
Compound types per group
Age group
Aging
Rejuve−nating
Other
Percent of compounds
c
MOA Aging compounds
Topoisomerase inh.
CDK inh.
MDM inh.
IMPDH inh.
CHK inh.
PKC activator
DNA synthesis inh.
Microtubule inh.
RNA synthesis inh.
Protein synthesis inh.
Farnesyltransferase inh.
Ribonucleotide reductase inh.
Dihydrofolate reductase inh.
Steroidal progestin
IKK inh.
Thymidylate synthase inh.
Ribosomal protein inh.
RNA polymerase inh.
PARP inh.
Akt inh.
STAT inh.
Aurora kinase inh.
MTOR inh.
KIT inh.
P38 MAPK inh.
PI3K inh.
score_pa
d
MOA Rejuvenating compounds
HDAC inh.
MEK inh.
Cardiac glycoside
Proteasome inh.
GSK−3 inh.
PI3K inh.
score_pa
e
Top compounds per cell lines
Compounds
AMG−232
mitoxantrone
topotecan
pralatrexate
CGP−60474
AZD−3514
AZD−9291
PMA
panobinostat
PD−0325901
vorinostat
trichostatin−a
bortezomib
MG−132
NEU
HL60
SHSY5Y
NCIH2110
NPC
THP1
HCT116
HS578T
MCF7
XC.L100
VCAP
XC.R10
HME1
U2OS
HEPG2
MCF10A
A375
XC.L10
A549
PC3
HA1E
MDAMB231
Cell lines
√score_pv
Cell line status
normal
pool
tumor

Figure 4

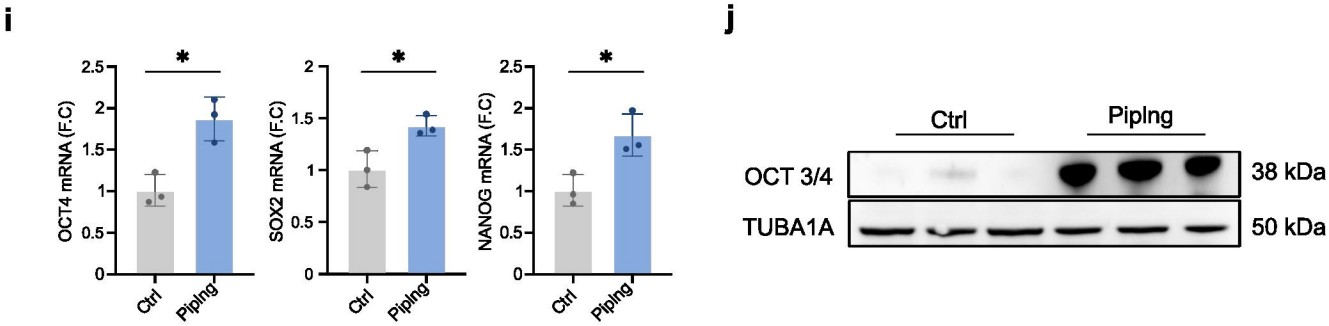
Senescence induction in responder A375



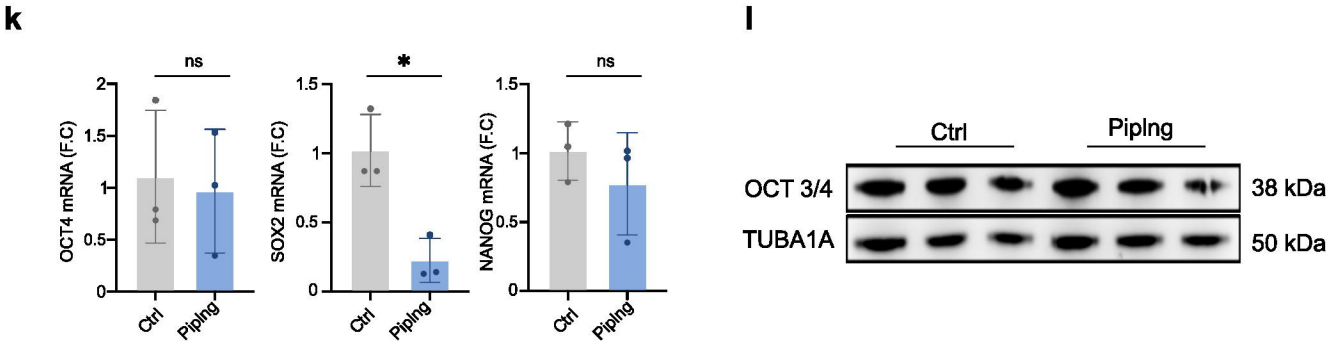
Senescence induction in non-responder MDA-MB-231



Stemness induction in responder PC3



Stemness induction in non-responder MCF7



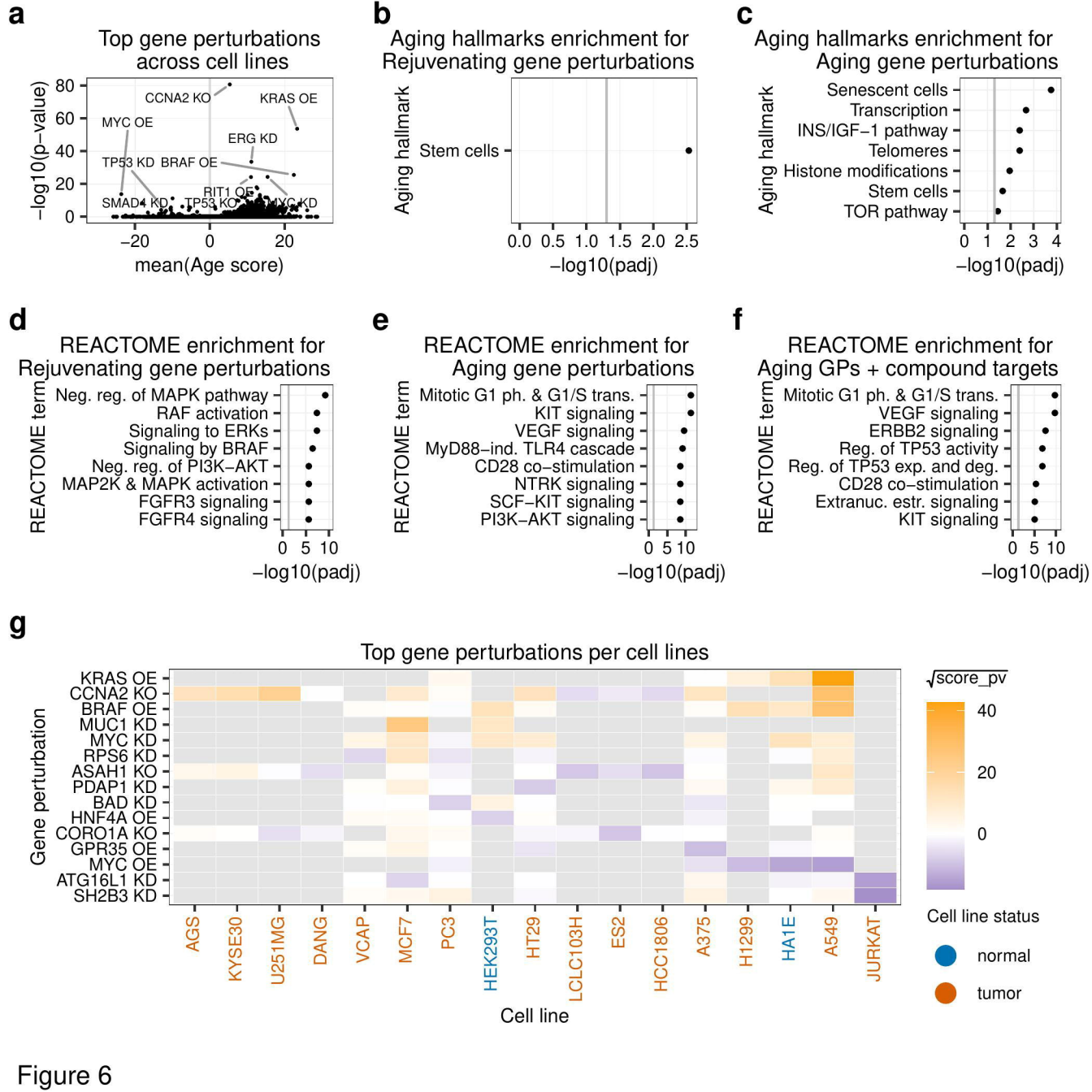

a
Top gene perturbations across cell lines
−log10(p−value)
mean(Age score)
CCNA2 KO
KRAS OE
MYC OE
ERG KD
TP53 KD
BRAF OE
RIT1 OE
SMAD4 KD
TP53 KO
MYC KD
b
Aging hallmarks enrichment for Rejuvenating gene perturbations
Aging hallmark
Stem cells
−log10(padj)
c
Aging hallmarks enrichment for Aging gene perturbations
Aging hallmark
Senescent cells
Transcription
INS/IGF−1 pathway
Telomeres
Histone modifications
Stem cells
TOR pathway
−log10(padj)
d
REACTOME enrichment for Rejuvenating gene perturbations
REACTOME term
Neg. reg. of MAPK pathway
RAF activation
Signaling to ERKs
Signaling by BRAF
Neg. reg. of PI3K−AKT
MAP2K & MAPK activation
FGFR3 signaling
FGFR4 signaling
−log10(padj)
e
REACTOME enrichment for Aging gene perturbations
REACTOME term
Mitotic G1 ph. & G1/S trans.
KIT signaling
VEGF signaling
MyD88−ind. TLR4 cascade
CD28 co−stimulation
NTRK signaling
SCF−KIT signaling
PI3K−AKT signaling
−log10(padj)
f
REACTOME enrichment for Aging GPs + compound targets
REACTOME term
Mitotic G1 ph. & G1/S trans.
VEGF signaling
ERBB2 signaling
Reg. of TP53 activity
Reg. of TP53 exp. and deg.
CD28 co−stimulation
Extranuc. estr. signaling
KIT signaling
−log10(padj)
g
Top gene perturbations per cell lines
Gene perturbation
KRAS OE
CCNA2 KO
BRAF OE
MUC1 KD
MYC KD
RPS6 KD
ASAH1 KO
PDAP1 KD
BAD KD
HNF4A OE
CORO1A KO
GPR35 OE
MYC OE
ATG16L1 KD
SH2B3 KD
AGS
KYSE30
U251MG
DANG
VCAP
MCF7
PC3
HEK293T
HT29
LCLC103H
ES2
HCC1806
A375
H1299
HA1E
A549
JURKAT
Cell line
√score_pv
Cell line status
normal
tumor

Figure 6

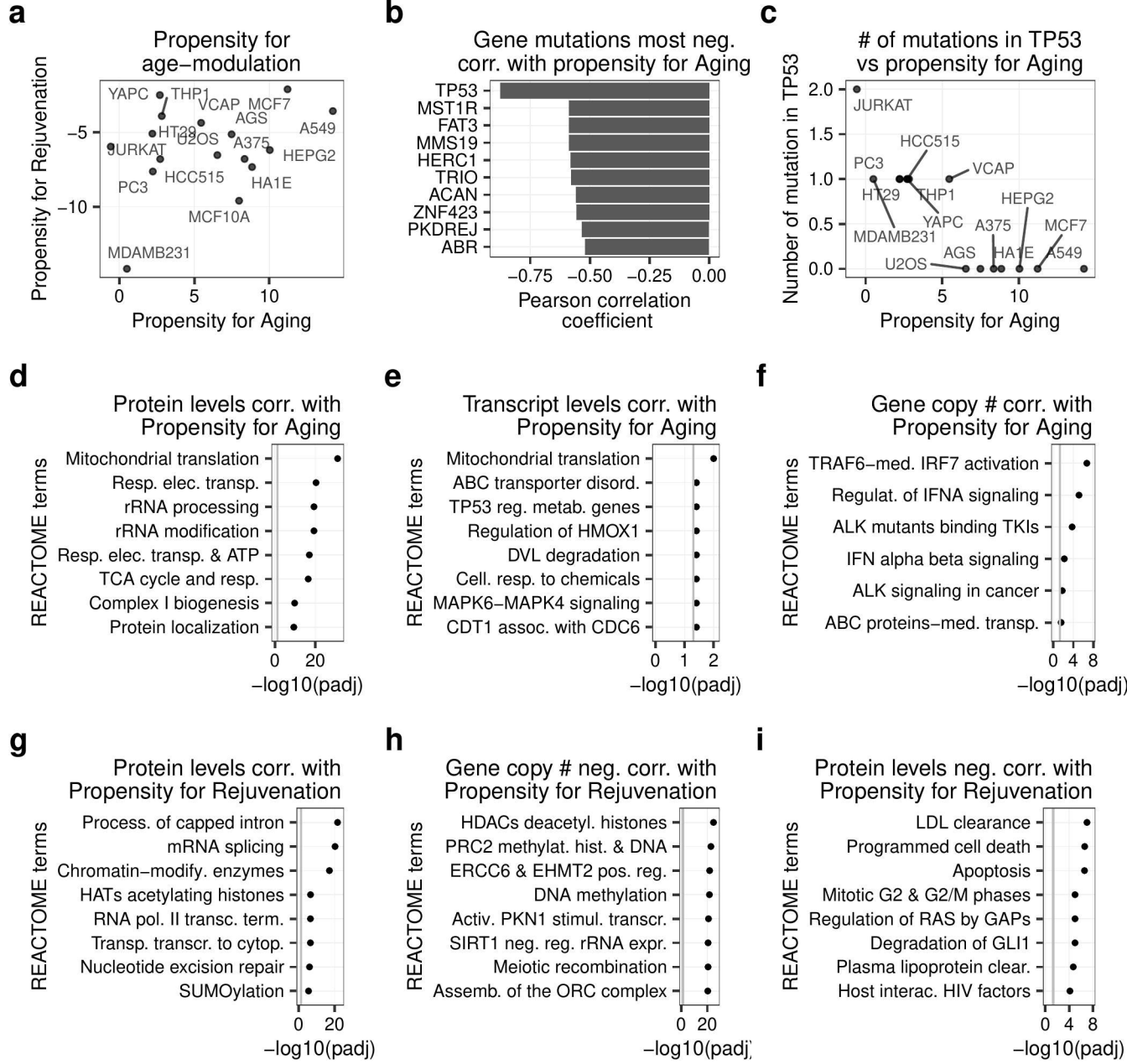

a
Propensity for age−modulation
Propensity for Rejuvenation
Propensity for Aging
YAPC
THP1
VCAP
MCF7
AGS
A549
HT29
U2OS
A375
JURKAT
HEPG2
HCC515
HA1E
PC3
MCF10A
MDAMB231
b
Gene mutations most neg. corr. with propensity for Aging
TP53
MST1R
FAT3
MMS19
HERC1
TRIO
ACAN
ZNF423
PKDREJ
ABR
Pearson correlation coefficient
c
# of mutations in TP53 vs propensity for Aging
Number of mutation in TP53
Propensity for Aging
JURKAT
HCC515
PC3
VCAP
HT29
THP1
HEPG2
YAPC
A375
MCF7
MDAMB231
AGS
HA1E
A549
U2OS
d
Protein levels corr. with Propensity for Aging
REACTOME terms
Mitochondrial translation
Resp. elec. transp.
rRNA processing
rRNA modification
Resp. elec. transp. & ATP
TCA cycle and resp.
Complex I biogenesis
Protein localization
−log10(padj)
e
Transcript levels corr. with Propensity for Aging
REACTOME terms
Mitochondrial translation
ABC transporter disord.
TP53 reg. metab. genes
Regulation of HMOX1
DVL degradation
Cell. resp. to chemicals
MAPK6−MAPK4 signaling
CDT1 assoc. with CDC6
−log10(padj)
f
Gene copy # corr. with Propensity for Aging
REACTOME terms
TRAF6−med. IRF7 activation
Regulat. of IFNA signaling
ALK mutants binding TKIs
IFN alpha beta signaling
ALK signaling in cancer
ABC proteins−med. transp.
−log10(padj)
g
Protein levels corr. with Propensity for Rejuvenation
REACTOME terms
Process. of capped intron
mRNA splicing
Chromatin−modify. enzymes
HATs acetylating histones
RNA pol. II transc. term.
Transp. transcr. to cytop.
Nucleotide excision repair
SUMOylation
−log10(padj)
h
Gene copy # neg. corr. with Propensity for Rejuvenation
REACTOME terms
HDACs deacetyl. histones
PRC2 methylat. hist. & DNA
ERCC6 & EHMT2 pos. reg.
DNA methylation
Activ. PKN1 stimul. transcr.
SIRT1 neg. reg. rRNA expr.
Meiotic recombination
Assemb. of the ORC complex
−log10(padj)
i
Protein levels neg. corr. with Propensity for Rejuvenation
REACTOME terms
LDL clearance
Programmed cell death
Apoptosis
Mitotic G2 & G2/M phases
Regulation of RAS by GAPs
Degradation of GLI1
Plasma lipoprotein clear.
Host interac. HIV factors
−log10(padj)

Figure 7