Title

Pasta, an age-shift 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, Stockholm, Sweden
2. Department of Molecular Biosciences, The Wenner-Gren Institute, Stockholm University, Stockholm, Sweden
3. Department of Cell and Molecular Biology, Karolinska Institute, Stockholm, 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

As the prevalence of age-related diseases rises, understanding and modulating the aging process is becoming a priority. Transcriptomic aging clocks (TACs) hold great promise for this endeavor, yet most are hampered by platform or tissue specificity and limited accessibility. Here, we introduce Pasta, a robust and broadly applicable TAC based on a novel age-shift learning strategy. Pasta accurately predicts relative age from bulk, single-cell, and microarray data, capturing senescent and stem-like cellular states through signatures enriched in p53 and DNA damage response pathways. Its predictions correlate with tumor grade and patient survival, underscoring clinical relevance. Applied to the CMAP L1000 dataset, Pasta identified known and novel age-modulatory compounds and genetic perturbations, and highlighted mitochondrial translation and mRNA splicing as key determinants of the cellular propensity for aging and rejuvenation, respectively. Supporting Pasta's predictive power, we validated pralatrexate as a potent senescence inducer and piperlongumine as a rejuvenating agent. Strikingly, chemotherapy drugs were highly enriched among pro-aging hits. Taken together, Pasta represents a powerful and generalizable tool for aging research and therapeutic discovery, distributed as an easy-to-use R package on GitHub.

Introduction

The growing prevalence of chronic non-communicable diseases that occur predominantly later in life have 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 interfere with aging[3]. Unlike chronological aging, which is inevitably proceeding at a set rate, 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 emerged to quantify biological age and estimate aging rates. These models draw on 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 gained prominence due to their high precision in estimating chronological and biological age, strong predictive value for morbidity and mortality, and the stability of DNA methylation (DNAm) changes, which enable the capture of cumulative aging effects[6]. However, DNAm clocks 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/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. Furthermore, the number of transcriptomic perturbation datasets, such as Perturb-Seq datasets[14] or the Tahoe-100 dataset[15], is rapidly increasing. Applying TACs to such datasets could expose the causal regulators of aging and rejuvenation with high translational potential: compounds that raise the age of cancer cells may boost therapy[16], whereas agents that lower normal-cell age bear regenerative promise, such as improvement in stem cell fitness[17], iPSC protocols[18], or partial reprogramming cocktails[19]. One study matched three tissue signatures to the CMAP L1000 perturbation atlas[13]—a large scale (> 3 million samples, >250 cell lines) bead-based (L1000 technology) genetic and chemical perturbation dataset—and validated novel modulators[20], but its narrow sample base (109 aging samples, two CMAP cell lines, only chemical perturbations) limited scope.

Human TACs have been developed in blood[21], muscle[22], fibroblasts[23], skin[24], and in multi-tissue settings[25–29]. More recently, single-cell TACs have been constructed using blood[30], immune cells[31,32], and PBMCs[33]. Despite this progress, no TAC has achieved widespread adoption comparable to epigenetic clocks. This is due to several limitations, including unreliable performance, lack of user-friendly software, and platform-specific constraints. Gene expression's sensitivity to environmental changes allows detection of transient states but also introduces greater variability and noise than DNA methylation data, demanding robust modeling across multiple datasets. Most TACs—including many single-tissue models and the multi-tissue RNAAgeCalc[26]—are trained on single datasets, leading to overfitting and poor generalization. Indeed, MultiTIMER[27], which was trained on multiple datasets, largely

outperform RNAAgeCalc[26] despite being trained on less samples (~3,000 vs ~9,600). Another key limitation is the lack of accessible tools. DeepQA and Shokhirev & Johnson's TACs offer no usable software[25,28], and RAPToR requires users to generate custom references without default human datasets[29]. Moreover, most multi-tissue TACs[25–28] necessitate steps like raw data reprocessing, batch correction, and model retraining, limiting their use to bulk RNA-Seq data and hampering reproducibility.

To address these limitations, we present Pasta (Predicting age-shift from transcriptomic analysis): a multi-platform, multi-tissue, ready-to-use TAC. Across diverse datasets it outperformed MultiTIMER[27], largely via p53-centric gene and pathway features. Pasta's predictions effectively distinguished between senescent, quiescent, and stem cells, in static and dynamic settings, indicating that Pasta is optimally suited to capture bidirectional cell state transitions from stemness to senescence. Pasta was a better predictor of tumor grade than chronological age, reflecting its usefulness for biological age estimation. Applied to CMAP L1000[13], Pasta revealed new age modulators—including pralatrexate, a potent senescence inducer—and revealed molecular traits dictating cell-line responsiveness to pro- or anti-aging perturbations. Pasta thus delivers a robust, interpretable framework for cross-platform biological-age estimation and systematic discovery of genetic or chemical modulators of aging and rejuvenation, with translational promise in oncology and regenerative medicine.

Results

Building a general-purpose transcriptomic aging clock using heterogeneous multi-tissue data

We aimed to create a versatile multi-tissue transcriptomic aging clock usable on any gene expression sample. Heterogeneous raw or normalized matrices from the Genotype-Tissue Expression Project (GTEx)[34], the Gene Expression Omnibus (GEO)[35], and the Expression Atlas[36] were rank-transformed, yielding 17,212 healthy profiles from 21 studies (18 bulk RNA-seq, 3 microarrays; Fig 1a, Table S1). We then compared three different modeling strategies: regression models (Fig. 1b), young vs old classifiers (Fig. 1c), and age-shift classifiers (Fig. 1d). Our proposed age-shift classifier mirrors twin neural network regression[37,38] and pairwise difference regression[39] or classification[40]: it trains on matched-tissue sample pairs separated by a defined age gap. Classifiers' predictions were converted to age scores (see Methods).

Using a leave-one-dataset out (LODO) strategy, we found that the 40-years age-shift model had the highest median Pearson correlation coefficient (PCC) when considering all samples (Fig. 1e) or only samples from young and old individuals (Fig. S1a), with accurate predictions across tissues and platforms (Fig S1b-S3). Models trained on un-ranked counts collapsed (median PCC < 0.1), confirming the need for rank transformation (Fig 1e, S1a). Next, we tested the best three models on single-cell RNA-Seq datasets. As only bulk transcriptomic data were used to train the models, we created pseudobulks by aggregating single cells using chunk sizes from 1 (single-cell) to 1024 to see how increased data reliability improved model performance. We found that the age-shift model performed best overall in all three datasets analyzed (Fig. 1f-h)[41–43]. Interestingly, the young vs old classifier failed in datasets without

samples in both the young and old categories (Fig. 1f,h), highlighting issues with fixed age cutoffs (Fig. 1c). Overall, we found that the 40-years age-shift model had the best performance for predicting relative age from heterogeneous transcriptomic data and we refer to it as Pasta in the remainder of this manuscript.

Pasta outperforms alternative approaches and leverages p53-related pathways to predict transcriptomic age

To evaluate Pasta's performance, we compared it with MultiTIMER, the current state-of-the-art, available, multi-tissue RNA-Seq aging clock[27]. We found that Pasta outperformed MultiTIMER in 6 out of 8 datasets from our LODO analysis (Fig. 2a, Table S2) and in 5 out of 6 independent datasets (Fig. 2b). These datasets represented diverse tissues (Table S2) which validate Pasta's multi-tissue capabilities.

We then conducted gene set enrichment analysis (GSEA) on the REACTOME, KEGG, miRDB, and CGP subsets[44] using Pasta's coefficients. The p53 tumor suppressor pathway was significantly enriched in the positive coefficients of the model in the KEGG, and CGP gene sets (Fig. 2c, Table S3). Additionally, irradiation stress terms, which are tightly linked to p53 dynamics[45], were also enriched. The strongest negative enrichments were targets of miR29, a p53 activator[46] which drives aging phenotypes in mice[47]. Interestingly, no pathway was significantly enriched when using the baseline regression model, indicating that the age-shift learning strategy amplified aging-related signal.

Individual coefficients mirrored these trends. RANBP17, carried the largest negative weight, nearly twice that of any other gene (Fig. 2d). Interestingly, RANBP17 has been shown to promote cell proliferation[48] and was proposed to be a master regulator of aging which decreases with age[49]. The strongest positive weights belonged to PCDHB3, a colorectal-cancer tumour-suppressor[50], CDKN2A (p16), a canonical p53 target and senescence marker (Fig. 2e), and ZMAT3 another direct p53 regulator.

In conclusion, Pasta operates as multi-tissue, and multi-platform relative age prediction model which grounds its predictions in tumor-suppressor genes (TP53, PCDHB3) and core ageing modulators (TP53, miR29, RANBP17).

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

Since Pasta bases its predictions on p53, a key regulator of cell cycle and senescence, we investigated whether it could be used to identify senescent cells. Pasta effectively separated senescent from proliferative cells in 19 out of 30 RNA-Seq datasets (AUC-ROC of 1) and yielded moderate discrimination (AUC-ROC > 0.7) in five additional datasets (Fig. 3a, Table S4). It likewise robustly distinguished senescent from quiescent cells in 4/5 datasets (Fig. 3b). In a 450-day serial-passage microarray experiment of primary fibroblasts from a 5-year-old donor until replicative senescence, Pasta's age scores tracked culture duration almost linearly (PCC = 0.896, Fig. 3c). We next asked whether Pasta can detect chemical inducers of senescence in single-cell data. Merging two studies that assayed the same 138 compounds in

A549 cells—one by scRNA-seq[51], the other by flow cytometry using a senescence reporter[52]—we observed that Pasta's age scores predicted senescence-inducing molecules, with AUC > 0.8 at the two highest doses (Fig. 3d).

Given Pasta's ability to identify very old cells, we next tested its sensitivity for detecting very young cells. Also here, Pasta effectively separated embryonic stem cells from fibroblasts in four bulk RNA-seq datasets (Fig. 3e) and induced Pluripotent Stem Cells (iPSCs) from fibroblasts in five more (Fig. 3f). During OSKM-mediated fibroblast reprogramming, Pasta's age scores declined linearly over time (PCC = -0.905, Fig. 3g). Conversely, Pasta's age scores had strong correlation with time in a differentiation experiment of human embryonic stem cells into liver cells (PCC = 0.926, Fig. 3h).

Taking all these observations together, we conclude that Pasta can be used to track cellular age in static and dynamic settings.

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

Studies using transcriptomic stemness prediction models have shown that higher stemness correlates with advanced tumor grade and poorer overall survival in cancer patients[53,54]. Recently, older cancer cell age has also been linked to adverse outcomes[16]. Indeed, senescent cells, both normal and cancerous, accumulate in certain tumors, can drive systemic inflammation, facilitate relapse, and promote immune evasion[16,55]. However, The Cancer Genome Atlas (TCGA) transcriptomic data have, to our knowledge, not yet been used to predict older cellular age, leaving it unclear whether poor prognosis is driven by high stemness alone or also by senescence. We addressed this gap by leveraging Pasta's capacity to identify cells spanning both youthful/high-stemness and aged/senescent-like states.

In line with previous results, we found that six out of twelve cancer types had a significantly negative correlation between age scores and tumor grade (Fig. 3i, j, Table S5). Interestingly, two cancer types, kidney renal clear cell carcinoma and pancreatic adenocarcinoma, had a significantly positive correlation between tumor age and tumor grade (Fig. 3i, k) suggesting that both increased and decreased biological age can be associated with tumor grade. While Pasta was trained to predict age, its predictions exhibited stronger correlation with tumor grade than donor age itself, strongly suggesting that Pasta's age scores reflect cellular biological age instead of merely chronological age (Fig. 3i).

We trained Cox proportional hazards regression models to examine associations between age or Pasta's age scores and all-cause mortality across 33 cancer types. After multiple testing correction, younger age scores were significantly linked to poorer survival in ACC, SARC, and SKCM (Fig. 3l,m; Table S6), while older age scores were similarly associated in THYM and LAML (Fig. 3l,n). Several other cancers showed strong unadjusted associations. In contrast, chronological age predicted worse survival in 18 cancer types and was not protective in any. Thus, both younger and older cancer cell age scores were linked to tumor grade and survival, suggesting Pasta scores may help stratify patient risk in a cancer-specific manner.

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

To systematically identify genetic and chemical perturbations influencing cellular age, we applied Pasta to the CMAP L1000 dataset[13], leveraging its compatibility with heterogeneous transcriptomic data. Age scores were plate-normalized using respective negative controls (DMSO for compounds; empty vectors for genetic perturbations). We identified 259 and 59 compounds significantly associated to older and younger age scores—termed 'Aging' and 'Rejuvenating' compounds, respectively (Fig. 4a; Table S7). Top Rejuvenating compounds included pluripotency-promoting molecules, such as RepSox and PD0325901[56]. Top Aging compounds included senescence inducers such as mitoxantrone, gemcitabine, and doxorubicin[57]. To evaluate Pasta's ability to identify compounds with strong age-modulatory effect, we annotated compounds using drug classes promoting senescence[57] and pluripotency[56]. Notably, 18.9% of Aging compounds belonged to senescence-inducing drug classes, versus <0.2% in the Other/Rejuvenating sets (Fig. 4b), while 16.9% of Rejuvenating compounds belonged to reprogramming-inducing drug classes, again versus <0.2% in the other groups (Fig. 4c). Furthermore, 27.1% of Rejuvenating compounds were annotated as both, compared to 1.54% and 0.07% in the Aging and Other groups (Fig. 4d), with all such hits being HDAC inhibitors—agents known to promote reprogramming[58,59] and also induce senescence[60–62]. When repeating the analysis using the base regression model (Fig. 1), only 12.7%/6.5%/19.4% (vs 18.9%/16.9%/27.1%) of Aging/Rejuvenating/Both sets showed these annotations (Fig. S4), confirming that Pasta's age-shift model yields more biologically informative predictions.

We next explored mechanisms underlying these compound effects using the mechanisms of actions (MOAs) annotated in the dataset. Over 26 MOAs were significantly enriched among Aging compounds (Fig. 4e; Table S8), including several pro-senescence pathways such as topoisomerase, CDK, MDM, and IMPDH inhibitors, largely related to growth, cell cycle, and signaling. Only six MOAs were enriched in Rejuvenating compounds (Fig. 4f), with HDAC inhibitors showing the strongest enrichment, consistent with their potent role in accelerating the reprogramming process[58,59]. This set also included MEK and GSK-3 inhibitors, common in reprogramming cocktails[56], as well as cardiac glycosides and proteasome inhibitors—both described as senolytics[63]. Thus, selective elimination of older cells in culture may lead to net rejuvenation (see Discussion).

Analyzing compound effects across cell lines, proteasome inhibitors MG-132 and bortezomib showed consistent rejuvenating effects, while mitoxantrone showed consistently strong aging effects (Fig. 4g; Table S9). Some HDAC inhibitors (e.g., trichostatin-A, vorinostat, panobinostat) were rejuvenating in most lines but aging in neuronal lines (NEU, NPC, SH-SY5Y), consistent with trichostatin-A's cytotoxicity in SH-SY5Y[64]. These findings support the dual role of HDAC inhibitors in promoting reprogramming and senescence[58–62].

To test if age-modulatory compounds affected cancer cell viability, we integrated Cancer Dependency Map (DepMap)[65] data. Both Aging and Rejuvenating sets were broadly toxic to cancer cells, with median dependency scores of –1.86 and –2.97, respectively (Fig. 4h). MOA analyses revealed a significant link between age and dependency scores only for MDM and IMPDH inhibitors (Fig. 4i,k; Table S10). Among all compounds, only two MDM inhibitors—AMG-232 and CGM-097—showed statistically significant effects (Fig. 4j,k; Table S11). These findings indicate that MDM inhibitors may kill cancer cells by accelerating cellular

aging, in line with the dual role of the p53–MDM2 axis in promoting both senescence and apoptosis[66].

Identification of potential cancer-specific chemotherapeutic compounds

Chemotherapeutic drugs are thought to exert their effects by increasing the biological age of both cancer and normal cells[16], thereby reducing tumor stemness and aggressiveness. Leveraging the scale of our dataset, we systematically tested this hypothesis. We examined cancers with ≥5 approved chemotherapeutics to assess enrichment within our three age-association groups (Aging, Rejuvenating, Other) across cell lines. Notably, 17 of 27 cancers were significantly enriched in the Aging group (Figure 5a; Table S12,13). Of the ten non-significant cancers, eight had few approved drugs and adjusted p-values <0.2. Only one cancer showed enrichment in the Rejuvenating group, and none in the Other group. These results not only support the pro-aging mechanism of chemotherapeutics[16] but also highlight our method's capacity to identify candidate chemotherapeutics in a data-driven manner—particularly valuable for poorly treatable cancers. Future work could expand this approach by using only cell lines relevant for a given cancer with validation in primary clinical material.

We also searched for compounds inducing stronger senescence in cancer cells than in normal cells, aiming to discover agents with potentially reduced side effects. Pralatrexate, a dihydrofolate reductase inhibitor, showed the strongest pro-aging effect—four times higher in cancer cell lines (Fig. 5b-c; Table S14). Additional promising compounds included Exifone (an HDAC1 activator) and Brequinar (a dihydroorotate dehydrogenase inhibitor), representing potential candidates for selective, low-toxicity chemotherapies.

Experimental validation of predicted small molecules' age-modulatory effects

To experimentally assess the predictive power of Pasta in identifying compounds capable of promoting age-modulatory effects in a cell type–dependent manner, we exposed melanoma A375 cells (responders; Table S15) and breast cancer MDA-MB-231 cells (non-responders) to pralatrexate, the 4th top Aging drug (Tables S7). In line with Pasta's prediction, we observed that pralatrexate induced cell enlargement and positivity for the Senescence-Associated β-Galactosidase (SA-β-Gal) assay (Figure 6a), transcriptional upregulation of CDKN1A (also known as p21) and the SASP component IL6 (Figure 6b), proliferative arrest marked by reduced EdU incorporation (Figure 6c), and downregulation of LMNB1 alongside upregulation of p21 protein levels (Figure 6d)—all consistent with the induction of a *bona fide* senescence response. Conversely, in MDA-MB-231 cells, pralatrexate induced a stress response lacking key hallmarks of cellular senescence, as defined by minimal morphological changes yet detectable SA-β-Gal positivity (Figure 6e), negligible variations in CDKN1A and IL6 mRNA levels (Figure 6f), increased proliferation and EdU incorporation (Figure 6g), and modest changes in LMNB1 accompanied by upregulation of p21 protein levels (Figure 6h).

We also examined the capacity of Piperlongumine—the 20th top rejuvenating agent according to Pasta scoring (Table S7)—to promote stemness features in prostate cancer PC3 cells (responders; Table S15). We found that Piperlongumine induced transcriptional upregulation

of the pluripotency transcription factors OCT4, SOX2, and NANOG, accompanied by increased Oct4 protein levels, in PC3 cells (Figure 6i,j).

These findings underscore the robust predictive power of Pasta in capturing cell type–specific age-modulatory effects, including the ability to distinguish age-neutral stress response from bona fide senescence. These data strongly support the potential application of Pasta for in-silico identification of senescence- or reprogramming-inducing compounds, including through the mining of existing perturbation datasets.

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

We next sought to define the genetic determinants of cellular age from the CMAP L100 dataset, including gene knockdown, knockout, and overexpression. Across cell lines, we identified 841 gene perturbations (GPs) significantly associated with older age scores ('Aging' GPs) and 54 associated with younger scores ('Rejuvenating' GPs). Top aging GPs included knockout of CCNA2/cyclin and knockdown of the oncogenes ERG and MYC (Fig. 7a; Table S16), consistent with previous observations[67–69]. Overexpression of KRAS, BRAF, and RIT1 also led to older scores, reflecting the tumor-suppressive mechanism of oncogene-induced senescence (OIS)[70]. Conversely, MYC overexpression and SH2B3 or SMAD4 knockdown produced younger scores, aligning with earlier reports[71–73]. Both TP53 knockout and knockdown were among the top 51 youngest GPs (out of 10,479), consistent with its central tumor-suppressive role and its mutation in 50-60% of human cancers[74].

To characterize these GPs, we performed pathway enrichment using the Open Genes[75] and REACTOME[76] databases (Fig. 7b–f; S4; Tables S16–S18), inverting overexpression effects to pool GPs results. The only Aging Hallmark[75] enriched in Rejuvenating GPs was "stem cells exhaustion" (Fig. 7b), while "senescent cells accumulation" was the most enriched in Aging GPs (Fig. 7c). Additional enrichments in Aging GPs were related to transcriptional regulation, insulin/IGF-1 signaling, and telomeres. In REACTOME, Rejuvenating GPs were enriched in RAF/MEK/ERK and FGFR signaling (Fig. 7d), 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. 7e). Integration with compound targets revealed Aging GP enrichment in p53-dependent cell cycle regulation and growth factor signaling (e.g., VEGF, ERBB2, KIT, extranuclear estrogen) (Fig. 7f; S4g).

Next, we identified GPs with cell-line-specific age effects. Some GPs had consistent age effect across multiple cell lines: MYC overexpression had a consistent rejuvenating impact, while MYC knockdown, CCNA2 knockout, and BRAF/KRAS overexpression drove aging (Fig. 7g; Table S19). Some GPs, like SH2B3 knockdown and ASAH1 knockout, showed opposing effects depending on cell line, highlighting context-dependent roles.

Finally, we assessed whether age-altering GPs affected cancer cell viability. Aging knockout GPs had negative DepMap dependency scores (median: -1), while other GP sets hovered around zero (Fig. 7h). REACTOME-based analysis identified five pathways where age scores had significant association with dependency (Fig. 7i; Table S20), including retinoic acid biosynthesis (Fig. 7j), OIS (Fig. 7k), G1/S DNA damage checkpoints, and TP53 stabilization.

These findings suggest that age-promoting GPs, particularly related to senescence and TP53, also simultaneously promote cancer cell death.

Defining the molecular determinants of cellular aging and rejuvenation

Next, we leveraged the comprehensive nature of the CMAP L1000 and DepMap datasets to define the molecular determinants of cellular propensity for age changes. Aging and rejuvenation propensities were computed for 19 cell lines by averaging age scores of significant age-regulatory perturbations across cell lines (see Methods). Results revealed marked heterogeneity across lines (Fig. 8a), with MDA-MB-231 showing the strongest rejuvenation propensity (-14.2 vs. -9.6 in MCF10A), consistent with its role as a model for aggressive, triple-negative breast cancer[77].

Correlating aging propensity with gene mutation burden, TP53 emerged as the only significantly associated gene (PCC = 0.875; Fig. 8b; Table S21). Cell lines with two TP53 mutations had negative aging-propensity scores (i.e., JURKAT), those with a single mutation scored below 6, and TP53-wild-type lines scored above 6 (Fig. 8c). These results are in line with the key role of TP53 as a tumor suppressor that induces senescence in highly proliferative cells[78,79] and suggests novel potential tumor suppressor genes (Fig. 8b). Consistent with this idea, top significant genes in Figure 8b included the known tumor suppressor genes FAT3[80], HERC1[81], and ABR[82,83].

To identify pathways influencing aging and rejuvenation propensities, we performed REACTOME-based enrichment on molecules (proteins, transcripts, gene copy numbers) correlated (|PCC| > 0.4) with each trait (Fig. 8d–i; S5; Table S22). Proteins whose abundance correlated with aging propensity were enriched in ‘mitochondrial translation’, ‘respiratory electron transport’, and ‘RNA processing’ (Fig. 8d). Similarly, aging propensity-associated transcripts were enriched in ‘mitochondrial translation’, ‘ABC transporter disorders’, and ‘TP53-regulates metabolic genes’ (Fig. 8e), suggesting mitochondrial activity is a central regulator of aging. This echoes recent findings linking p53, mitochondrial translation, and the senescence-associated secretory phenotype (SASP; see Discussion)[84]. Copy number–based enrichments included interferon-related pathways—e.g., ‘TRAF6-mediated IRF7 activation’ and ‘IFNA signaling’ (Fig. 8f)—consistent with type I interferons' role in SASP[85], DNA damage–induced[86], and oncogene-induced senescence[87], as well as their pro-aging effects[88].

Rejuvenation propensity correlated with protein abundance in ‘mRNA splicing’, ‘HATs acetylate histones’, and ‘chromatin modifying enzymes’ (Fig. 8g). Notably, ‘HDACs deacetylate histones’ was the top negatively correlated term for rejuvenation-linked copy numbers (Fig. 8h), followed by PRC2-related and DNA methylation pathways. These findings underscore the influence of histone-modifying enzymes on reprogramming capacity, aligning with HDAC inhibitors’ roles in reprogramming[58,59] and the key role of PRC2 in pluripotency[89] and as an aging/rejuvenation biomarker[90–92]. Finally, proteins negatively correlated with rejuvenation propensity were enriched in apoptosis and checkpoint modules, including programmed cell death and G2/M control (Fig. 8i). This result is consistent with evasion of cell death being a hallmark of cancer[93] and p53/p21 acting as barrier to reprogramming[94].

Discussion

In this study, we introduced Pasta, a versatile multi-tissue transcriptomic aging clock that accurately predicts relative age across diverse transcriptomic datasets. We demonstrated its superior performance over conventional regression and classification approaches, validated its efficacy in both bulk and single-cell data, and showcased its utility in predicting cellular aging, senescence, and stemness, as well as in cataloging age-modulatory perturbations across cell lines. Collectively, these findings establish Pasta as a robust and generalizable tool for aging research and translational applications.

A key strength of Pasta lies in its adaptability to diverse transcriptomic platforms. Trained on raw and normalized, non-preprocessed RNA-Seq and microarray data, its heterogeneous training set enabled the model to learn age-predictive patterns based solely on the relative ordering of gene expression. Consequently, Pasta can be applied directly to bulk, single-cell, or L1000 data without realignment or batch correction. It outperformed MultiTIMER[27], the current state-of-the-art, available, multi-tissue TAC, on bulk RNA-Seq and is distributed as a user-friendly and readily-available R package (https://github.com/jsalignon/pasta), underscoring its value to the aging field.

Pasta employs an innovative age-shift learning approach based on pairwise sample training, conceptually related to twin neural networks[37,38] and pairwise regression/classification methods[39,40], known to improve prediction accuracy. Our analysis confirmed performance gains across bulk and single-cell datasets. Moreover, Pasta relied on more biologically relevant features (Fig. 2c) and generated more meaningful predictions on CMAP L1000 data (Fig. S4) than a conventional regression model, positioning age-shift as a powerful method for predicting relative age.

Strong evidence implicates DNA damage response (DDR), cell cycle, and p53 signaling as core biological drivers of Pasta's predictions. Enrichment analysis revealed p53- and irradiation-related pathways in Pasta's gene signatures (Fig. 2c), and key predictors like CDKN2A/p16 and ZMAT3 are p53 targets (Fig. 2e). Topoisomerase inhibitors, which induce DNA damage, and CDK inhibitors, which inhibit the cell cycle, were the two most significantly enriched MOAs for Aging compounds (Fig. 4e). Finally, G1/S phase transition and regulation of TP53 were among the most significantly enriched pathways for Aging perturbations with consistent chemical and genetic effect (Fig. 7f). DDR activation in response to genotoxic stress triggers p53-mediated senescence[95], and a recent study proposed DDR as a minimal unit of aging clocks[96], where DDR, cell cycle, and senescence represent aging over hours, days, and months. Within this framework, Pasta may capture transcriptomic aging at various timescales by leveraging predictive features from distinct temporal layers.

Pasta showed strong predictive power for identifying cellular senescence, stemness, and age-related dynamics such as reprogramming and differentiation. Notably, it achieved this without being trained on senescent or stem cell data. For example, the REO stemness index[54], a leading bulk transcriptomic stemness predictor, used GTEx data as a negative class during training (non-stem), whereas Pasta was trained directly on GTEx (among other datasets). This suggests that healthy donor transcriptomes might enhance stemness and senescence prediction models, helping to address the scarcity of senescent and stem cell training data. These

findings also raise questions about the true nature of epigenetic clocks' predictions; i.e., are they predicting biological age at the organismal (i.e., health status) or cellular (stemness/senescence) level? Supporting the later hypothesis, recent studies reported a partial overlap between aging and senescence-related CpGs[97] and the control of stemness by DNA methylation marks in astrocytes[98]. A quantitative assessment of how well epigenetic clocks predict stemness and senescence could help interpret results from such clocks.

Clinically, Pasta's age scores significantly correlated with tumor grades and survival outcomes across multiple cancer types (Fig. 3i-n), supporting the potential clinical utility of Pasta in stratifying patients in distinct risk groups. Furthermore, the fact that Pasta's age scores were more strongly associated with tumor grade than age itself (Fig. 3i) indicates that Pasta's age scores reflect biological rather than chronological age.

CMAP L1000 analysis showed marked enrichment of senescence-inducing perturbations in the Aging group (Fig. 4b, 7c) and of reprogramming- and stemness-associated perturbations in the Rejuvenating group (Fig. 4c,d, 7b). Notably, Aging perturbations (259 compounds, 841 GPs) greatly outnumbered Rejuvenating ones (59 compounds, 54 GPs), underscoring the diverse triggers of accelerated aging versus the precision required for rejuvenation. Perturbation effects were largely consistent across both cancer and normal cell lines (Tables S9, S19), indicating conservation of aging and rejuvenation pathways in health and disease.

A recent study showed that OSKM expression in cancer cells can induce senescence and apoptosis[99]. Furthermore, our top Rejuvenating compounds—MG-132 and bortezomib—have been reported to impair OSKM-mediated reprogramming and reduce stem cell pluripotency[100]. These discrepancies may be due to CMAP's short exposure time of ~24h for most samples. Indeed, proteasome inhibitors like MG-101 and MG-132 initially upregulate Oct4 (day 2) in adipose precursor cells, but suppress it at later stages (days 25–30)[101]. Thus, our results likely reflect the early stages of reprogramming, which may be consistent across cancer and normal cells. Therefore, they have particular relevance for partial reprogramming, aimed at restoring cell function with minimal identity change, to limit cancer risk[102].

Unexpectedly, proteasome inhibitors and cardiac glycosides were among the top Rejuvenating compounds (Fig. 4f). Interestingly, the cardiac glycoside ouabain, ranked 34th (Table S7), was recently predicted to exert rejuvenating effects in mice using a transcriptomic mortality clock[103]. CT-200783 and piperlongumine—structurally related—were the only top-20 Rejuvenating compounds outside HDAC, MEK, or proteasome inhibitor classes. Notably, proteasome inhibitors, cardiac glycosides, piperlongumine, and HDAC inhibitors have been described as having senolytics properties[63]. Interestingly, different studies reported either senolytic or pro-senescence effects of bortezomib[63], indicating that cellular response may depend on cell state or environmental conditions, such as time of compound exposure. Our results suggest that the observed rejuvenation may result from selective removal of older cells, thereby rejuvenating the population as a whole. Furthermore, the surviving cells could feed from the remnants of dead cells to further boost growth and stemness[104,105]. Supporting this, HDAC inhibitors have shown rejuvenating properties[106], and proteasome inhibitors can initially induce OCT4[101]. Future studies should dissect the relative roles of old cell clearance, growth from dead cells debris, and intrinsic rejuvenation in these compounds' rejuvenating effects.

Among the chemical perturbations identified, pralatrexate emerged as a novel senescence-inducing compound. Ranked fourth overall for senescence induction (Table S7) and validated experimentally (Fig. 6), it aligns with the known effects of other dihydrofolate inhibitors like methotrexate[107–109] (ranked 98th), pemetrexed[110], and pyrimethamine[111]. Notably, pralatrexate ranked first for inducing senescence more strongly in cancer cells (Fig. 5b–c; Table S14), suggesting it may be an alternative to methotrexate for limiting tumor growth with reduced side effects[109]. FDA-approved in 2009 for relapsed or refractory peripheral T-cell lymphoma[112], its senescence-inducing effects had not been previously reported. This underscores the power of transcriptomic aging clocks like Pasta to uncover unrecognized therapeutic properties of known drugs and may help clarify pralatrexate's risks and benefits, particularly by leveraging insights into the role of senescence in normal and cancer cells.

Integrating CMAP L1000 with DepMap data revealed key molecular determinants of aging and rejuvenation. Transcriptomic and proteomic analyses showed that cell responses to pro-senescence stimuli were strongly influenced by mitochondrial translation activity (Fig. 8d,e). This aligns with extensive evidence linking mitochondrial metabolism, ROS production, and senescence[113]. Mitochondrial translation, essential for biogenesis, generates ROS[114], which are required for p21-mediated senescence[115]. Additionally, p53 has been shown to enhance mitochondrial translation in senescent cells, supporting cytokine production via SASP[84]. Genetically, cells with higher type I interferon gene copy numbers were more prone to aging responses (Fig. 8f), consistent with the pro-senescence role of interferon signaling in SASP and neighboring cell effects[116]. Together, these findings support a model in which enhanced mitochondrial translation and interferon signaling prime cells for pro-senescence responses via ROS and inflammatory pathways.

Conversely, rejuvenation was primarily governed by epigenetic factors (Fig. 8g,h), consistent with known reprogramming barriers[106,117]. Notably, reduced copy numbers in PRC2 and HDAC genes increased sensitivity to rejuvenating stimuli (Fig. 8h), in line with PRC2's role as a universal biomarker of aging and rejuvenation[90] and HDAC inhibitors' efficacy in reprogramming[106]. Altogether, our analysis identifies both established and novel regulators of aging and rejuvenation, offering mechanistic insights and potential therapeutic targets.

In conclusion, Pasta is a versatile, robust, and biologically informative transcriptomic aging clock with broad utility—from fundamental research on aging, senescence, and stem cells to clinical diagnostics and therapeutic discovery (Figure S7). By enhancing the accuracy and generalizability of transcriptomic aging clocks, we enabled systematic mining of genetic and chemical perturbations to identify causal regulators of cellular age. This approach holds strong translational promise, facilitating data-driven identification of pro-aging perturbations for cancer therapy and rejuvenating interventions for regenerative medicine and future age-reversal strategies.

Methods

Construction of the aging transcriptomic database

GTEx Protected Access Data version 8 (17,382 samples) was used to obtain precise age information[34]. 1,674 GTEx samples with a RIN score below 6 were discarded. 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[118] for GEO[35] datasets and using the package ExpressionAtlas[36] for Expression Atlas[119] datasets. The biomaRt R package[120] 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 had a total of 17,212 samples, with 14,571 from the GTEx dataset, and contained 8,113 genes in common, which were also included in the landmark or best inferred genes sets in the CMAP L1000 dataset[13], and which 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. Finally, 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. Finally, 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[121], and using the following parameters: *s_lamba = '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 12 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[27]. 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.

Conducting gene set enrichment analysis

GSEA was performed using the GSEA function from the *clusterProfiler* R package[122]. Gene sets were retrieved via the msigdbr function from the *msigdbr* package[123], aggregating sets from the CPG, 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

The supplementary Table 1 from the SenCID manuscript[124] 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 samples out of 603 which could be downloaded and analyzed. For stem cell datasets, samples were manually annotated. Data was 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[125].

Analyzing TCGA data

Histological tumor grades were obtained from Ping et al[126] for samples from BRCA patients and from Liu et al[127] for samples from patients with other cancer types, as described in Zheng et al[54]. Only cancer types with histological tumor grades ranging from G1 and G4 were retained, excluding BLCA samples. The TCGAbiolinks R package was used to download gene expression data and obtain age and overall survival information[128]. 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[129]. 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 actions (MOAs) were manually annotated for 60 compounds. Missing compounds 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 the Table 1 of Petrova et al[57]. Reprogramming-inducing compounds were obtained from the Supplementary Table 1 of Knyazer et al[56]. 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 mechanism of actions (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 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 50 perturbations were selected for each group, which ensured to display 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 OpenGene[75]. REACTOME gene sets were obtained from the Human MSigDB Collections[44], using the msigdbr R package[123]. Only REACTOME pathways with 300 genes or fewer were retained (1,565 out of 1,615). Over-representation analysis was performed using the *enricher* function from the clusterProfiler R package[122] with the parameter *universe* set to all the unique genes in the input database.

Analyzing CMAP L1000 data – Analysis of anti-cancer drugs

Anti-cancers drugs were obtained from the CancerDrugs_DB[130] (build date: 28/11/24). We selected drugs approved from the FDA or the EMA or Euro nationally approved drugs which kept 311 out of 321 drugs. We then kept 132 drugs that were present in the CMAP L1000 dataset. For each cancer, over-representation of its anti-cancer 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[131]. Cell lines 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 (Mechanism of actions for compounds and REACTOME pathways for CRISPR) were tested by fitting linear model 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 0 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 within both the Rejuvenation and the 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. Out of 1,724 genes, only TP53 showed a significant association. PCCs between cell line's 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 Institutet, Sweden). MDA-MB-231 (human breast cancer) cells were provided by Pr. Kirsty Spalding (Karolinska Institutet, Sweden). PC3 cells were kindly provided by Dr. Yvonne Ceder (Lund University). Jurakt 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 experiment. 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 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) (Biorad,#11704156) 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 (Sigma-Aldrich, #10735078001, w:v in TBS), followed by an overnight incubation with primary antibodies specific for Waf1/Cip1/CDKN1A p21 (Santa Cruz Biotechnology, #sc-6246), 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) RT-qPCR. 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 sdetection 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 Table S1. 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 Pasta R package, including pre-trained models, functions for data download, processing, and age prediction, is freely 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 Institutet (KID2016-00207 and KID2021-00495) and by an ICMC project grant 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), Longevity Impetus Grant from Norn Group, 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 Institutet 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 how to validate our transcriptomic aging clock. We thank Bader Zarrouki (AstraZeneca) for scientific discussions.

Conflicts of interest

The authors declare no conflict of 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., 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. 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.

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

19. 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.

20. 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.

21. 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.

22. 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*.

23. 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.

24. 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.

25. Shokhirev, M.N., and Johnson, A.A. Modeling the human aging transcriptome across tissues, health status, and sex. Aging Cell *n/a*, e13280. https://doi.org/10.1111/acel.13280.

26. 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.

27. 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.

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

29. 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.

30. 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.

31. 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.

32. 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.

33. 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.

34. 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.

35. 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.

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

37. 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.

38. 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.

39. 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.

40. 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.

41. 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.

42. 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.

43. 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.

44. 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.

45. 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.

46. 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.

47. 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.

48. 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.

49. 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.

50. 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.

51. 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.

52. 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.

53. 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.

54. 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.

55. Park, S.S., Choi, Y.W., Kim, J.-H., Kim, H.S., and Park, T.J. (2021). Senescent tumor cells: an overlooked adversary in the battle against cancer. Exp Mol Med *53*, 1834–1841. https://doi.org/10.1038/s12276-021-00717-5.

56. 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.

57. 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.

58. 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.

59. 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.

60. 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.

61. 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.

62. 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.

63. 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.

64. 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.

65. Tsherniak, A., Vazquez, F., Montgomery, P.G., Weir, B.A., Kryukov, G., Cowley, G.S., Gill, S., Harrington, W.F., Pantel, S., Krill-Burger, J.M., et al. (2017). Defining a Cancer Dependency Map. Cell *170*, 564-576.e16. https://doi.org/10.1016/j.cell.2017.06.010.

66. Huang, Y., Che, X., Wang, P.W., and Qu, X. (2024). p53/MDM2 signaling pathway in aging, senescence and tumorigenesis. Seminars in Cancer Biology *101*, 44–57. https://doi.org/10.1016/j.semcancer.2024.05.001.

67. 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.

68. 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.

69. 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.

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. 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.

72. 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.

73. 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.

74. 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.

75. 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.

76. 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.

77. 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.

78. 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.

79. 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.

80. 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.

81. 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.

82. 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.

83. 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.

84. 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.

85. 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.

86. 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.

87. 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.

88. 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.

89. 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.

90. 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.

91. 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.

92. 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.

93. Seidel, E., and von Karstedt, S. (2022). Extrinsic cell death pathway plasticity: a driver of clonal evolution in cancer? Cell Death Discov. *8*, 1–6. https://doi.org/10.1038/s41420-022-01251-7.

94. 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.

95. 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.

96. Han, J.-D.J. (2024). The ticking of aging clocks. Trends in Endocrinology & Metabolism *35*, 11–22. https://doi.org/10.1016/j.tem.2023.09.007.

97. Wagner, W. (2019). The Link Between Epigenetic Clocks for Aging and Senescence. Front Genet *10*, 303. https://doi.org/10.3389/fgene.2019.00303.

98. Kremer, L.P.M., Cerrizuela, S., El-Sammak, H., Al Shukairi, M.E., Ellinger, T., Straub, J., Korkmaz, A., Volk, K., Brunken, J., Kleber, S., et al. (2024). DNA methylation controls stemness of astrocytes in health and ischaemia. Nature *634*, 415–423. https://doi.org/10.1038/s41586-024-07898-9.

99. Pedrosa, P., Zhang, Z., Nuñez-Quintela, V., Macias, D., Ge, J., Denholm, M., Dyas, A., Estevez-Souto, V., Lado-Fernandez, P., Gonzalez, P., et al. (2024). Inhibition of lung tumorigenesis by transient reprogramming in cancer cells. Cell Death Dis *15*, 1–16. https://doi.org/10.1038/s41419-024-07207-2.

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

101. 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.

102. Singh, P.B., and Zhakupova, A. (2022). Age reprogramming: cell rejuvenation by partial reprogramming. Development *149*, dev200755. https://doi.org/10.1242/dev.200755.

103. 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.

104. Huang, Q., Li, F., Liu, X., Li, W., Shi, W., Liu, F.-F., O'Sullivan, B., He, Z., Peng, Y., Tan, A.-C., et al. (2011). Caspase 3–mediated stimulation of tumor cell repopulation during cancer radiotherapy. Nat Med *17*, 860–866. https://doi.org/10.1038/nm.2385.

105. Haak, V.M., Huang, S., and Panigrahy, D. (2021). Debris-stimulated tumor growth: a Pandora's box? Cancer Metastasis Rev *40*, 791–801. https://doi.org/10.1007/s10555-021-09998-8.

106. 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.

107. 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.

108. 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.

109. 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.

110. 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.

111. 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.

112. 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.

113. 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.

114. 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.

115. 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.

116. 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.

117. Ji, S., Xiong, M., Chen, H., Liu, Y., Zhou, L., Hong, Y., Wang, M., Wang, C., Fu, X., and Sun, X. (2023). Cellular rejuvenation: molecular mechanisms and potential therapeutic interventions for diseases. Sig Transduct Target Ther *8*, 1–39. https://doi.org/10.1038/s41392-023-01343-5.

118. 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.

119. 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.

120. 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.

121. 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.

122. 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.

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

124. 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.

125. 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.

126. 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.

127. 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.

128. 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.

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

130. 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.

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

132. 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.

133. 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.

134. 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 is a novel modeling approach that enables cross-platform aging clocks.

(**a**) Overview of the evaluated modelling strategies. Models were trained and tested exclusively on samples from healthy donors. (**b-d**) Schematic diagrams of the different modeling strategies. (**e**) Comparative performance in predicting relative age, assessed with leave-one-dataset-out cross-validation. All samples were rank-transformed except for REG_RC and AS40_RC where raw counts ('RC') 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 scRNA-Seq aging atlases of the retina[42] (**f**), skin[43] (**g**), and lung[41] (**h**). Abbreviations: REG, regression; AS30/40/50, age-shift classifier with 30/40/50 years of age difference, TC_30_70|TC_40_60, traditional classifiers with young cutoffs at 30/40 years and old cutoffs at 70/60 years, RC: raw counts.

Figure 2. Pasta is a state-of-the-art, multi-tissue relative age-prediction model whose output is predominantly driven by p53-linked signals.

(**a-b**) Pearson Correlation Coefficient (PCC) of age scores for Pasta and MultiTIMER vs true ages for healthy individuals. (**a**) Results from the leave-one-dataset-out analysis. (**b**) Results on external datasets. (**c**) Gene-Set Enrichment Analysis (GSEA) of the coefficients of the Pasta model, using the REACTOME, CPG, KEGG and MIRDB gene sets from the Human MSigDB Collections. (**d-e**) Top 10 most positive (**d**) and negative (**e**) coefficients.

Figure 3. Pasta identifies senescent and stem cells and stratifies 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), across multiple RNA-Seq datasets. (**c**) Pasta's predictions on a microarray senescence time course dataset[132]. (**d**) Pasta was applied to pseudobulks from a SciPlex dataset generated by exposing the A549[51] cell line to 138 drugs, and the resulting age scores were evaluated for their capacity to recover senescence-inducing drugs, as reported in Wang et al 2017[52]. (**e-f**) Pasta's ability to discriminate between stem cells (**e**) or induced-pluripotent stem cells (iPSC) (**f**) and fibroblasts, assessed by AUC-ROC. (**g**) Pasta's age scores on a bulk RNA-Seq reprogramming time course dataset[133]. (**h**) Pasta's age scores on a bulk RNA-Seq liver differentiation time course dataset[134]. (**i**) Correlation of patient 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 model 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**) Volcano plot showing the most significant Rejuvenating (negative age score) and Aging (positive age score) compounds. Points in yellow, blue, and purple indicate compounds from drug classes that induce senescence[57], promote reprogramming[56], or do both, respectively. For display purposes, p-values of zero were set to 10^{-300}. (**b-d**) Proportion of compounds annotated as senescence-inducing (**b**), pro-reprogramming (**c**), or both (**d**) for significantly age-increasing ('Aging'), age-decreasing ('Rejuvenating') compounds, and other compounds ('Other'). (**e-f**) All significantly enriched Mechanisms Of Actions (MOAs) for the Aging (**e**) and Rejuvenating (**f**) compounds. Colors are defined as in (a). The grey lines show the 0.05 adjusted p-value cutoff. (**g**) Heatmap of compounds by cell line with the most significant age scores (see Methods). Colors indicate the square root of a metric balancing significance and effect size: *-log10(adjusted p-value) * mean(age score)*. Missing values are shown in grey. (**h**) 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. (**i-j**) MOAs (**i**) and compounds (**j**) with the strongest association between age score and dependency. The yellow

and blue colors indicate a negative and positive Pearson correlation coefficient between age score and dependency, respectively. (**k**) Mean age score vs dependency for MDM inhibitors. Red and blue points show results for AMG-232 and CGM-097, respectively, across different cell lines.

Figure 5. Pasta can prioritize novel anti-cancer drugs with potentially diminished side-effects.

(**a**) Cancers whose anti-cancer drugs are significantly enriched in age-modulatory compounds across all cell lines, with yellow and blue colors indicating enrichment in the Aging and Rejuvenating sets, respectively. Only cancers with at least 5 approved anti-cancer drugs according to CancerDrugs_DB[130] were selected. 17 out 27 cancers were significant for the Aging group. (**b**) Top 10 compounds whose age score is 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).

Figure 6. Experimental validation of pralatrexate as a senescence-inducing compound and piperlongumine as a rejuvenating compound.

(**a**) Representative images from Senescence Associated Beta Galactosidase assay performed on A-375 cells exposed to Pralatrexate (Prltx). Doxorubicin (Doxo) was used as positive control for senescence induction. Scale Bar: 10 μm. (**b**) qPCR analysis of CDKN1A and IL6 mRNA levels in A-375 cells in response to Prtlx and Doxo.; N=3 biological replicates; *** $p< 0.001$ (one-way ANOVA). (**c**) Cell proliferation assay by Edu incorporation performed in A-375 cells exposed to Prltx and Doxo (one representative experiment with n=3 replicates/condition). *** $p< 0.001$ (one-way ANOVA, calculated on Edu^+ cells). (**d**) Immunoblotting showing LMNB1, CDKN1A and GAPDH protein levels in A-375 cells exposed to Prltx and Doxo (N=3 biological replicates). (**e**) Representative images from Senescence Associated Beta Galactosidase assay performed on MDA-231 cells exposed to Pralatrexate (Prltx). Doxorubicin (Doxo) was used as positive control for senescence induction. Scale Bar: 10 μm. (**f**) qPCR analysis of CDKN1A and IL6 mRNA levels in MDA-231 cells in response to Prltx and Doxo.; N=3 biological replicates; *$p< 0.05$ (one-way ANOVA). (**g**) Cell proliferation assay by Edu incorporation performed in MDA-231 cells exposed to Prltx and Doxo (one representative experiment with n=3 replicates/condition). ** $p< 0.01$; *** $p< 0.001$ (one-way ANOVA, calculated on Edu + cells). (**h**) Immunoblotting showing LMNB1, CDKN1A and GAPDH protein levels in A-375 cells exposed to Prltx and Doxo (N=3 biological replicates). (i) qPCR analysis of Oct4, Sox2 and Nanog levels in PC3 cells exposed to Piperlongumine (Piplng). * $p< 0.05$ Student's t test. (j) Immunoblotting showing OCT3/4 and alpha-tubulin protein levels in PC3 cells incubated with Piplng..

Figure 7. Pasta identifies age-modulatory gene perturbations.

(**a**) Volcano plot displaying the most significant Rejuvenating (negative age score) and Aging (positive age score) gene perturbations (GPs). (**b-f**) Enrichment analysis using the Hallmarks

of Aging[75] (**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 types of GPs (oe, sh, and xpr) were pooled, with overexpression results inverted (i.e., MYC_oe was included in the Aging GP set). Note that the Hallmarks of Aging gene sets are not directional (i.e., genes in the set "stem cell exhaustion" are related to stem cell in general; the last nouns were not used when compiling these gene sets). (**g**) Heatmap of GPs by cell lines with the most significant age scores (see Methods). Colors indicate the square root of a metric balancing significance and effect size: *-log10(adjusted p-value) * mean(age score)*. Missing values are shown in grey. (**h**) Boxplots and violin plots of dependency scores for GPs by cell line in the Aging, Rejuvenating and Other groups. The knockdown (KD) sets contain RNAi data for DepMap and shRNA data from CMAP, while the knockout (KO) sets contain CRISPR data from both databases. Number of entries per set: Other (KD: 27,243, KO: 46,684 entries), Aging (KD: 152, KO: 55), Rejuvenation (KD: 6, KO: 13). (**i**) REACTOME pathways whose PCCs are most significantly associated with dependency scores in the knockout set. All terms were negatively correlated with dependency scores. The grey line shows the 0.05 adjusted p-value cutoff. (**j**) Mean age score vs dependency for GPs in the retinoic acid (RA) biosynthesis pathway. Red and blue points show results for RDH13 and ALDH1A1, respectively, across different cell lines. (**k**) Mean age score vs dependency for GPs in the oncogene-induced senescence pathway. Red and blue points show results for MDM2 and TP53 knockout, respectively, across different cell lines. Abbreviations: oe, over-expression; sh, shRNA; xpr, CRISPR; KD, knockdown; KO, knockout.

Figure 8. Pasta uncovers the molecular determinants of cell lines' propensity for aging and rejuvenation.

(**a**) Propensity of various cell lines to become older or younger following exposure to age-regulatory perturbations (see Methods). (**b**) Gene mutations most strongly negatively correlated with the propensity for Aging. (**c**) Number of TP53 mutations versus the propensity of cells for Aging. (**d-i**) REACTOME pathway enrichment for genes whose protein levels (**d, g, i**), expression levels (**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**).

Supplementary Figures legends

Figure S1. 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 individauls. The x-axis shows the AUC-ROC for predicting whether the first sample is younger than the second in each evaluation pair.

Figure S2. 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.

Figure S3. 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.

Figure S4. Analysis of L1000 data with the baseline regression model recovers fewer known age-modulatory compounds.

Proportion of compounds annotated as senescence-inducing (**a**), pro-reprogramming (**b**), or both (**c**) for significantly age-increasing ('Aging'), age-decreasing ('Rejuvenating') compounds, and other compounds ('Other'). Analyses were performed identically to Figure 4b-d but employed the baseline regression model from Figure 1 rather than the age-shift model.

Figure S5. REACTOME enrichment analysis of 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) across all cell types that showed consistency between the pooled gene perturbations and the compound target analyses.

Figure S6. REACTOME enrichment analysis of pathways significantly associated with the 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**).

Figure S7. Graphical summary of the study.

Supplementary Tables legends

Table S1. Datasets used to train and evaluate the age models.

The *dataset* column includes GEO accession numbers, Expression Atlas IDs, or project names (GTEx and TCGA). The *n_samples* column indicates how many samples from healthy individuals were used to train and evaluate the models. The *n_pairs* column indicates how many sample pairs were used to train the final AS40/Pasta model. The *age_Q1|Q2|Q3* columns represent the first, second and third quartile of donor age.

Table S2. Benchmarking results comparing Pasta and MultiTIMER.

This table contains the Pearson Correlation Coefficients (PCCs) between true age and predicted age score for Pasta and MultiTIMER in the datasets shown in panels 2a and 2b, and the tissue type for each dataset.

Table S3. GSEA pathway enrichment results based on Pasta's coefficients.

Gene-Set Enrichment Analysis (GSEA) of the coefficients of the Pasta model, using the REACTOME, CPG, KEGG and MIRDB gene sets from the Human MSigDB Collections.

Table S4. AUC-ROC results for senescence and stem cell datasets.

This table reports the AUC-ROC results for datasets shown in panels 3a, 3d, 3e, and 3f. It includes the number of samples (or pseudobulks for panel 3d) for each group as well as the tissue type for each comparison, and compound dose for panel 3d.

Table S5. Association between age scores or age and tumor grade.

Results from correlation analyses between tumor grade and either chronological age or Pasta age scores. Abbreviations: padj – adjusted p-value, signif – significant result.

Table S6. Association between age scores, chronological age, and all-cause survival.

Results from Cox proportional-hazards regression analyses using scaled Pasta age predictions and scaled chronological ages. Abbreviations: padj – adjusted p-value.

Table S7. Compounds with significant age-associations across all cell types.

Abbreviations: padj – adjusted p-value, n – number of samples, n_cell – number of cell lines, med_age – median age scores.

Table S8. Mechanisms of actions (MOAs) with significant age association across all cell types.

bg indicates the background and lists the MOA of all compounds included in the analyses. Abbreviations: padj – adjusted p-value, tot – total.

Table S9. Compounds with significant age associations in specific cell types.

Abbreviations: padj – adjusted p-value, n – number of samples, med_age – median age score.

Table S10. MOAs with a significant association between age score and dependency.

This table contains the significance of the association between mean age score and dependency score across cell lines and compounds for each MOA, adjusted by compound effects. Only entries with an unadjusted p-value below 0.05 are included in this table. The column *n_cellXdrug* indicates the number of data points (i.e., compound-cell line pair) used to fit the regression model. Abbreviations: PCC – Pearson Correlation Coefficient, padj – adjusted p-value.

Table S11. Compounds with a significant association between age score and dependency.

This table contains the significance of the association between mean age score and dependency for each compound. Only entries with a non-adjusted p-value below 0.05 are included in this table. The column *n_cell* indicates the number of data points (i.e., cell lines) used to fit the regression model. Abbreviations: PCC – Pearson Correlation Coefficient, padj – adjusted p-value.

Table S12. Drugs used to treat cancers.

Drugs used to treat the various cancers analyzed in Figure 5. The *age_group* column indicates if a drug has a pro-aging effect, pro-rejuvenating effect, or no effect (group Other) across cell lines.

Table S13. Enrichment of anti-cancer drugs in age-modulatory compound groups.

Abbreviations: in_cat_yes/in_cat_no – number of drugs that are used to treat the indicated cancer that are in the given age-modulatory compound group (in_cat_yes) or that are not in

this group (in_cat_no), out_cat_yes/out_cat_no – number of drugs that are used to treat the indicated cancer that are in the other age-modulatory compound groups (out_cat_yes) or that are not in these groups (out_cat_no), padj – adjusted p-value.

Table S14. Cancer-specific pro-aging compounds.

Abbreviations: sd – standard deviation, n_cell – number of cell lines, n_rep – number of entries, padj – adjusted p-value, age_diff – difference in mean age effect between tumor and normal cell lines.

Table S15. Results for pralatrexate and piperlongumine in all cell types.

Abbreviations: padj – adjusted p-value, n – number of samples, med_age – median age score.

Table S16. Gene perturbations with significant age associations across all cell types.

Abbreviations: padj – adjusted p-value, n – number of samples, n_cell – number of cell lines, med_age – median age score.

Table S17. Hallmarks of Aging significantly enriched in Aging and Rejuvenating GPs.

Table S18. REACTOME terms significantly enriched in Aging and Rejuvenating GPs.

Table S19. Gene perturbations with significant age associations in specific cell types.

Abbreviations: padj – adjusted p-value, n – number of samples, med_age – median age score.

Table S20. REACTOME terms with a significant association between age score and dependency.

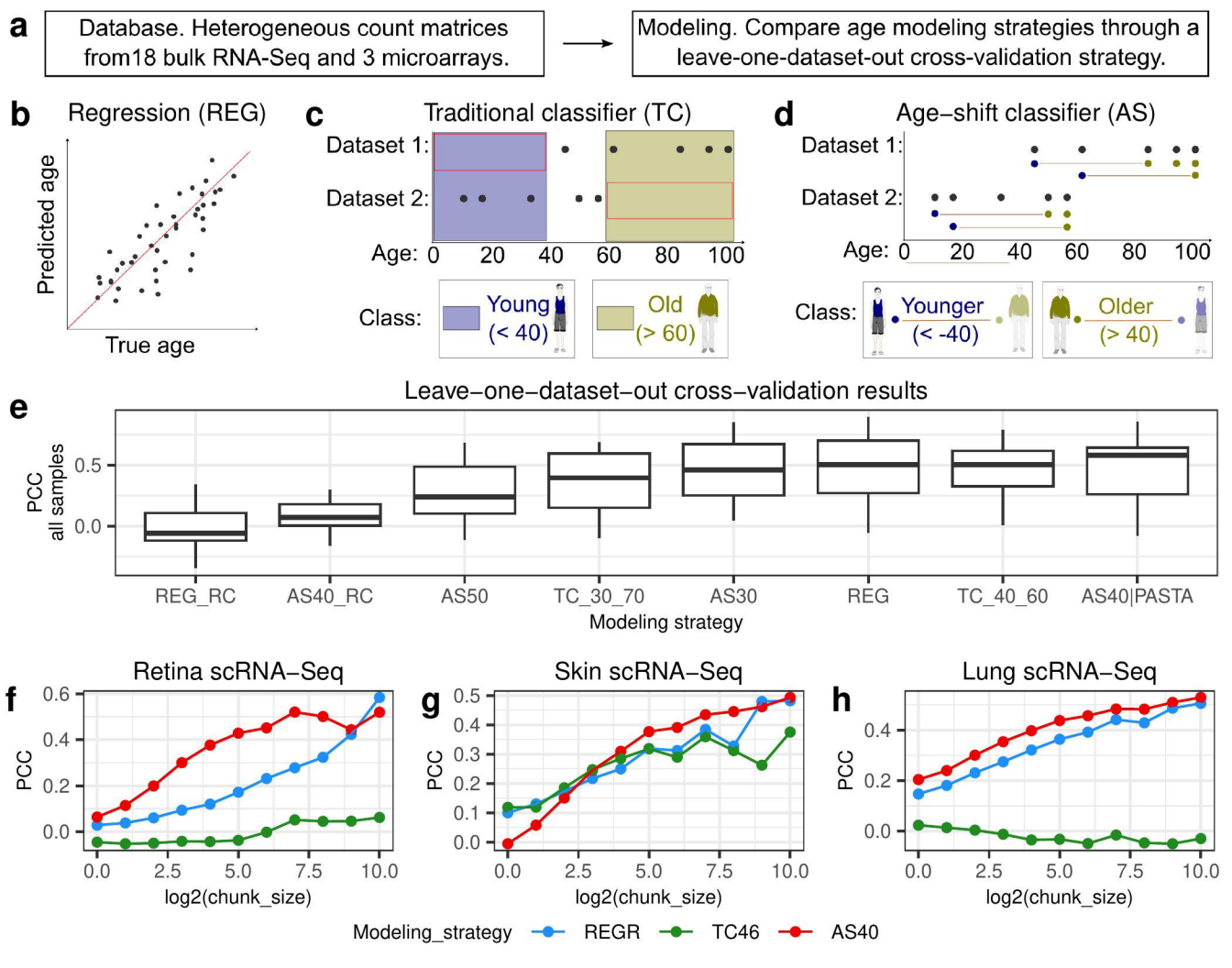
This table contains the significance of the association between mean age score and dependency score across cell lines and gene knockout for each REACTOME term, adjusted for the effect of gene knockouts. Only entries with a non-adjusted p-value below 0.05 are included in this table. The column *n_cellXgeneKO* indicates the number of data points (i.e., GP-cell line pair) used to fit the regression model. Abbreviations: PCC – Pearson Correlation Coefficient, padj – adjusted p-value, REAC_term – REACTOME term.

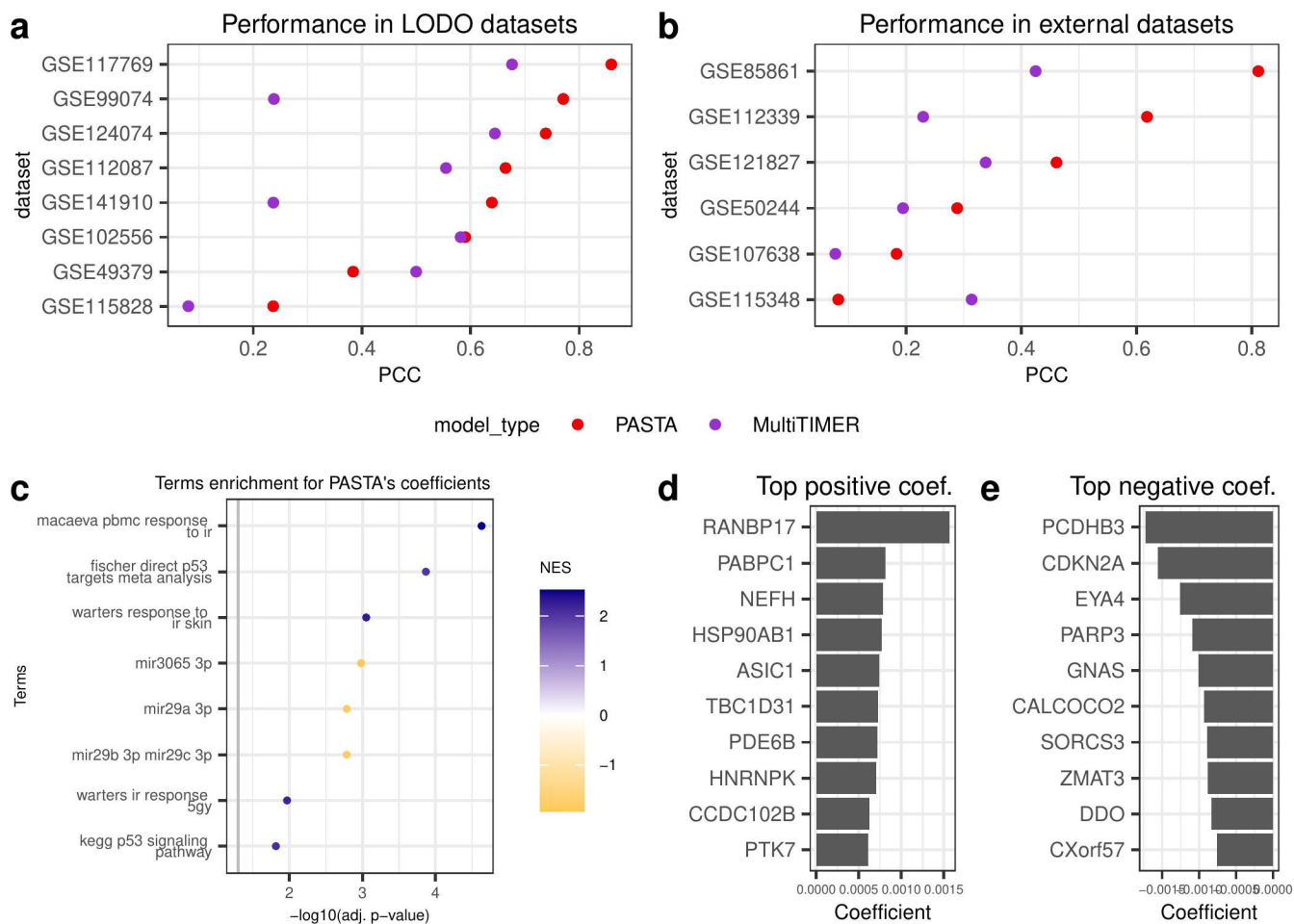
Table S21. Genes whose mutation count across cell lines is associated with cellular propensity for Aging.

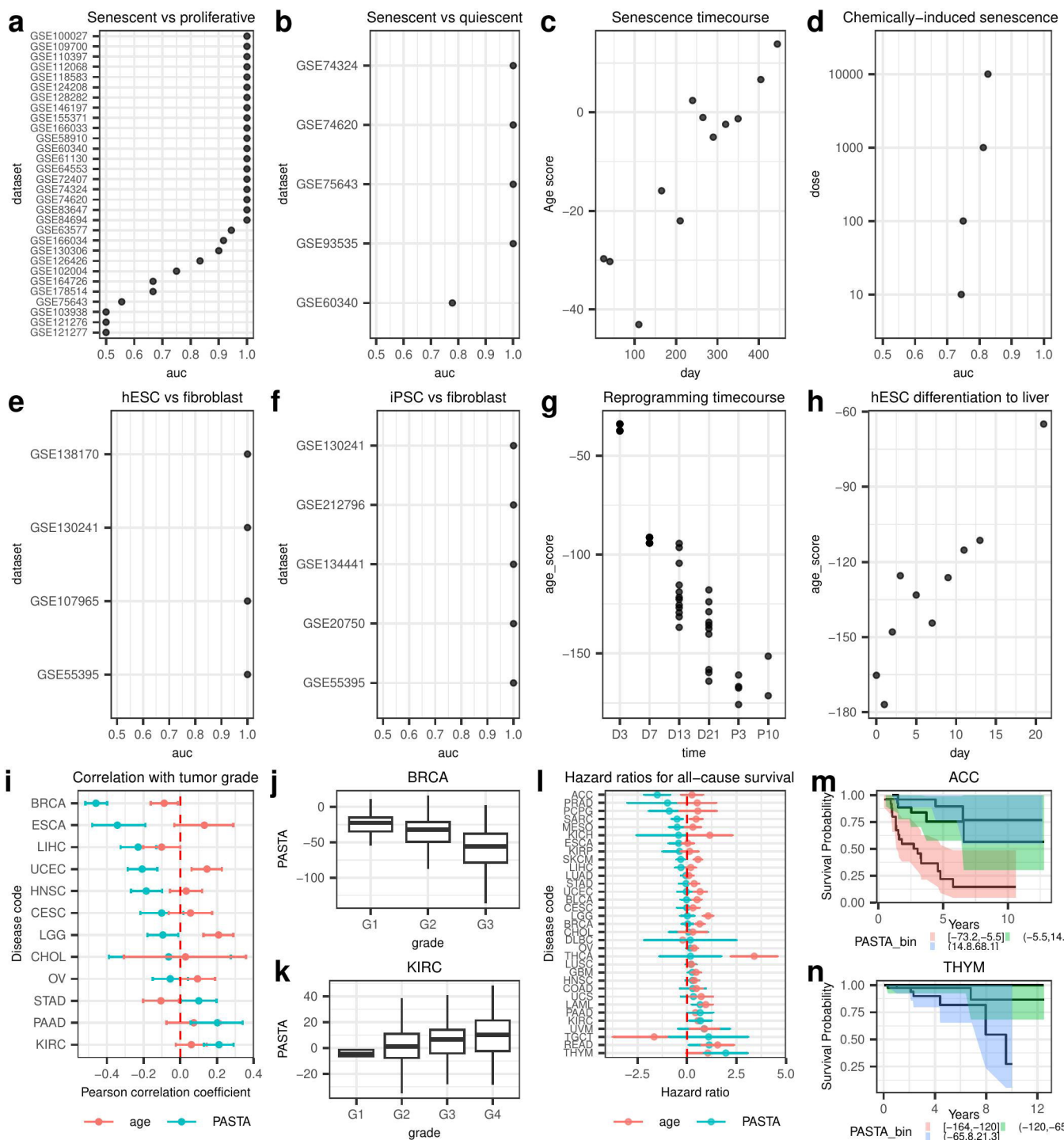
Only entries with a non-adjusted p-value below 0.05 are included in this table. Abbreviations: PCC – Pearson Correlation Coefficient, padj – adjusted p-value, n_cell – number of cell lines, n_mut – number of mutations, n_cell_mut – number of cell lines with at least one mutation.

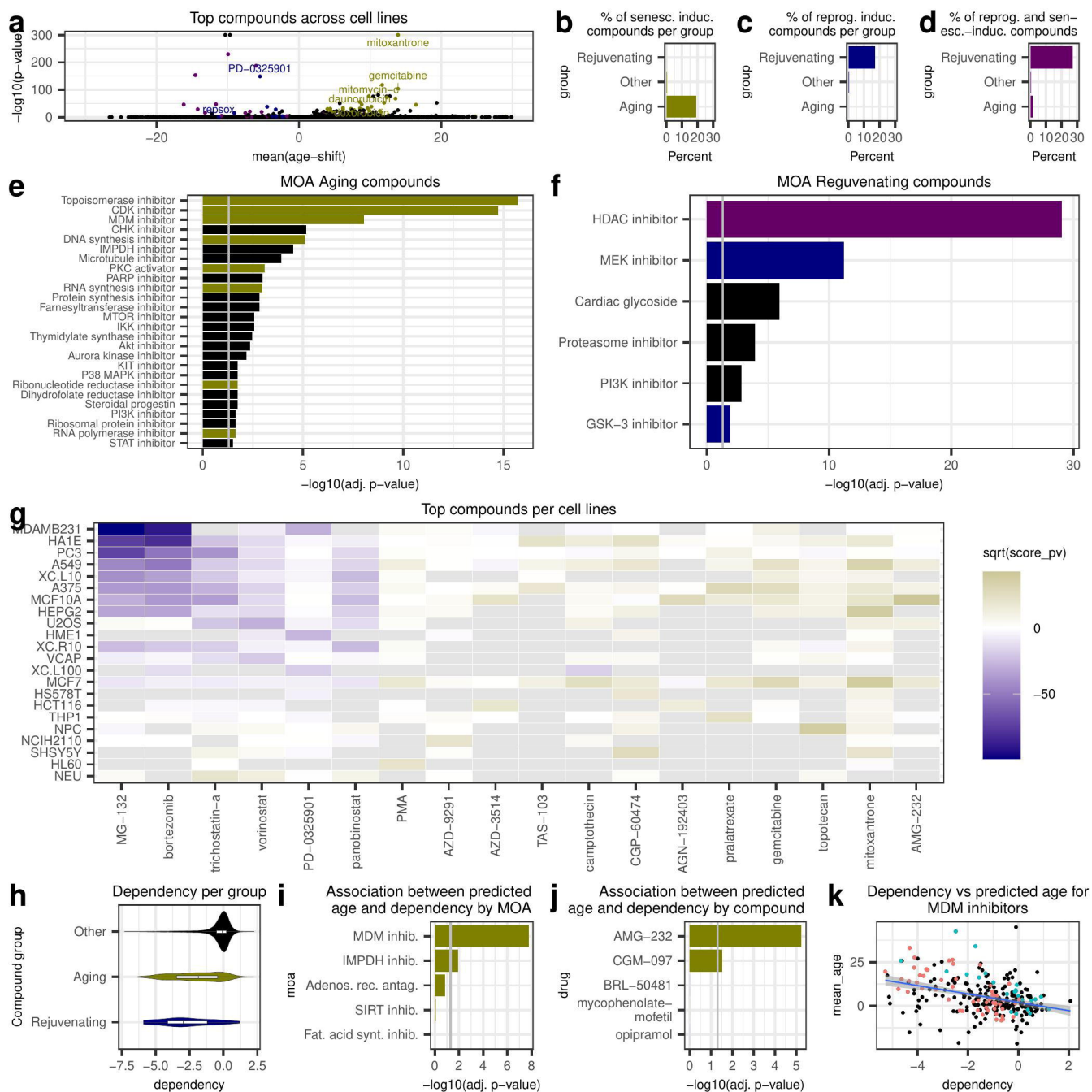
Table S22. REACTOME terms significantly associated with the propensity for Aging and Rejuvenation.

Abbreviations: PCC_dir – direction of the Pearson Correlation Coefficient.

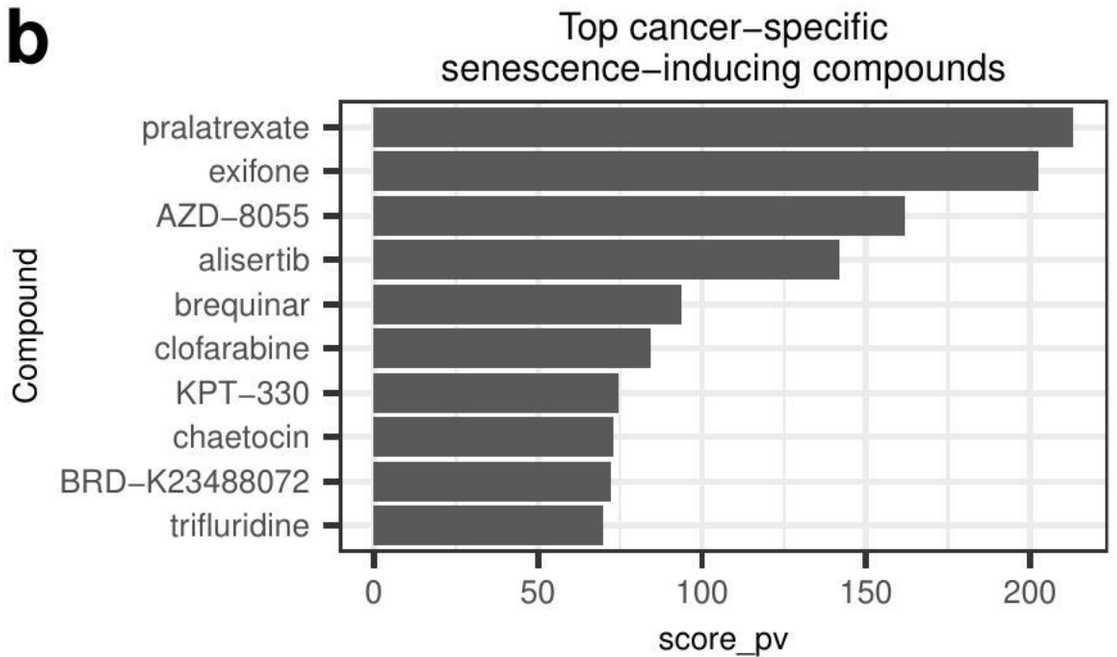

a
Database. Heterogeneous count matrices from18 bulk RNA-Seq and 3 microarrays.
Modeling. Compare age modeling strategies through a leave-one-dataset-out cross-validation strategy.
b
Regression (REG)
Predicted age
True age
c
Traditional classifier (TC)
Dataset 1:
Dataset 2:
Age: 0 20 40 60 80 100
Class:
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
PCC all samples
0.5
0.0
REG_RC
AS40_RC
AS50
TC_30_70
AS30
REG
TC_40_60
AS40|PASTA
Modeling strategy
f
Retina scRNA−Seq
PCC
log2(chunk_size)
g
Skin scRNA−Seq
PCC
log2(chunk_size)
h
Lung scRNA−Seq
PCC
log2(chunk_size)
Modeling_strategy
REGR
TC46
AS40

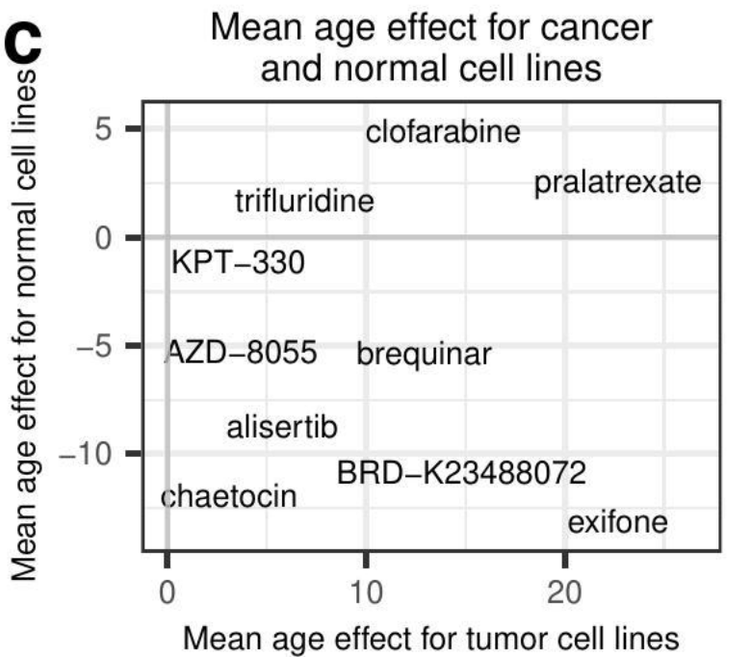

a
Performance in LODO datasets
GSE117769
GSE99074
GSE124074
GSE112087
GSE141910
GSE102556
GSE49379
GSE115828
dataset
0.2
0.4
0.6
0.8
PCC
b
Performance in external datasets
GSE85861
GSE112339
GSE121827
GSE50244
GSE107638
GSE115348
dataset
0.2
0.4
0.6
0.8
PCC
model_type
PASTA
MultiTIMER
c
Terms enrichment for PASTA's coefficients
macaeva pbmc response to ir
fischer direct p53 targets meta analysis
warters response to ir skin
mir3065 3p
mir29a 3p
mir29b 3p mir29c 3p
warters ir response 5gy
kegg p53 signaling pathway
Terms
2
3
4
−log10(adj. p−value)
NES
2
1
0
−1
d
Top positive coef.
RANBP17
PABPC1
NEFH
HSP90AB1
ASIC1
TBC1D31
PDE6B
HNRNPK
CCDC102B
PTK7
0.0000 0.0005 0.0010 0.0015
Coefficient
e
Top negative coef.
PCDHB3
CDKN2A
EYA4
PARP3
GNAS
CALCOCO2
SORCS3
ZMAT3
DDO
CXorf57
−0.0015 −0.0010 −0.0005 0.0000
Coefficient

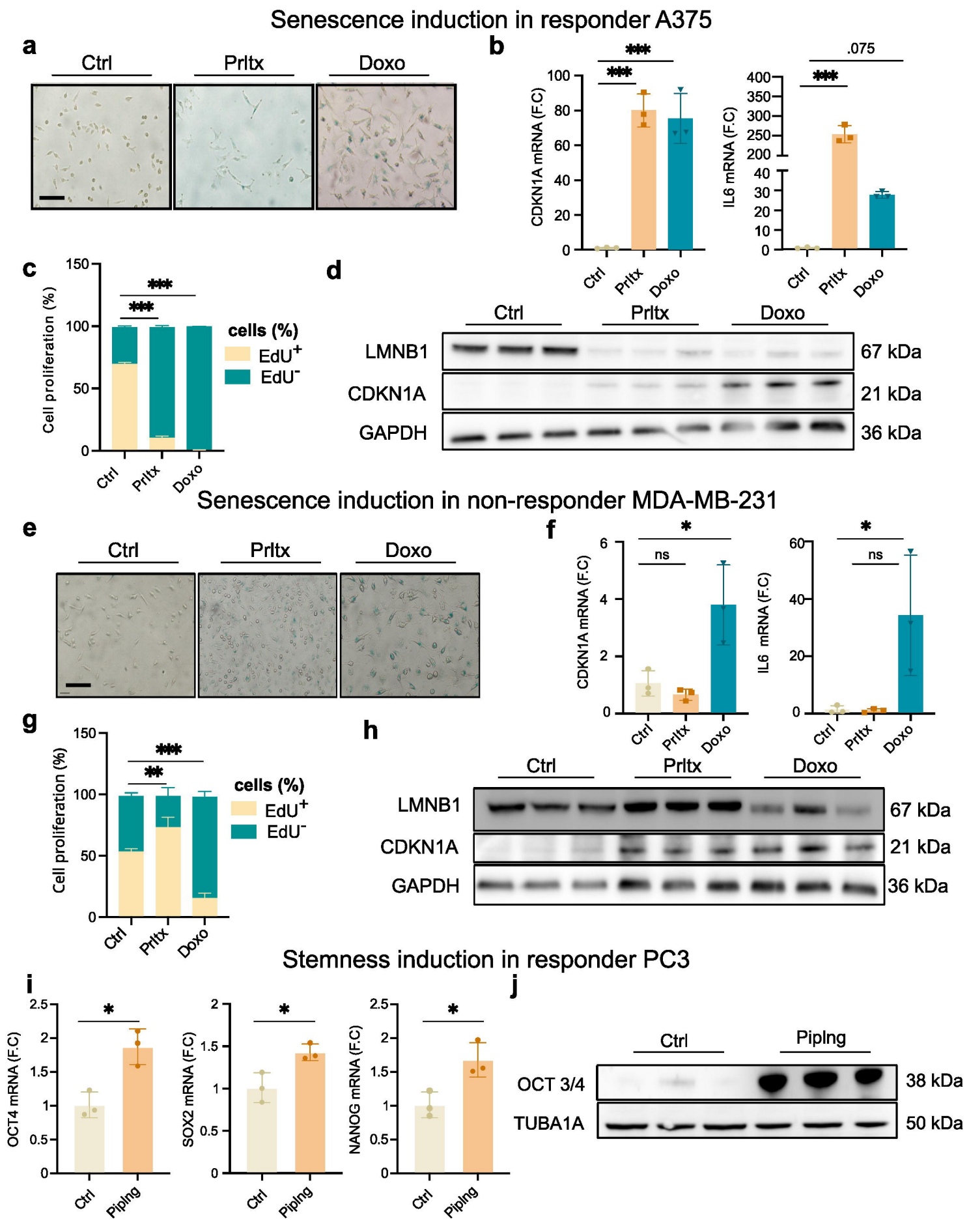

a
Senescent vs proliferative
dataset
auc
b
Senescent vs quiescent
GSE74324
GSE74620
GSE75643
GSE93535
GSE60340
c
Senescence timecourse
Age score
day
d
Chemically−induced senescence
dose
e
hESC vs fibroblast
GSE138170
GSE130241
GSE107965
GSE55395
f
iPSC vs fibroblast
GSE130241
GSE212796
GSE134441
GSE20750
GSE55395
g
Reprogramming timecourse
age_score
time
D3
D7
D13
D21
P3
P10
h
hESC differentiation to liver
day
i
Correlation with tumor grade
Disease code
BRCA
ESCA
LIHC
UCEC
HNSC
CESC
LGG
CHOL
OV
STAD
PAAD
KIRC
Pearson correlation coefficient
age
PASTA
j
BRCA
PASTA
grade
G1
G2
G3
k
KIRC
G4
l
Hazard ratios for all−cause survival
Hazard ratio
m
ACC
Survival Probability
Years
PASTA_bin
n
THYM

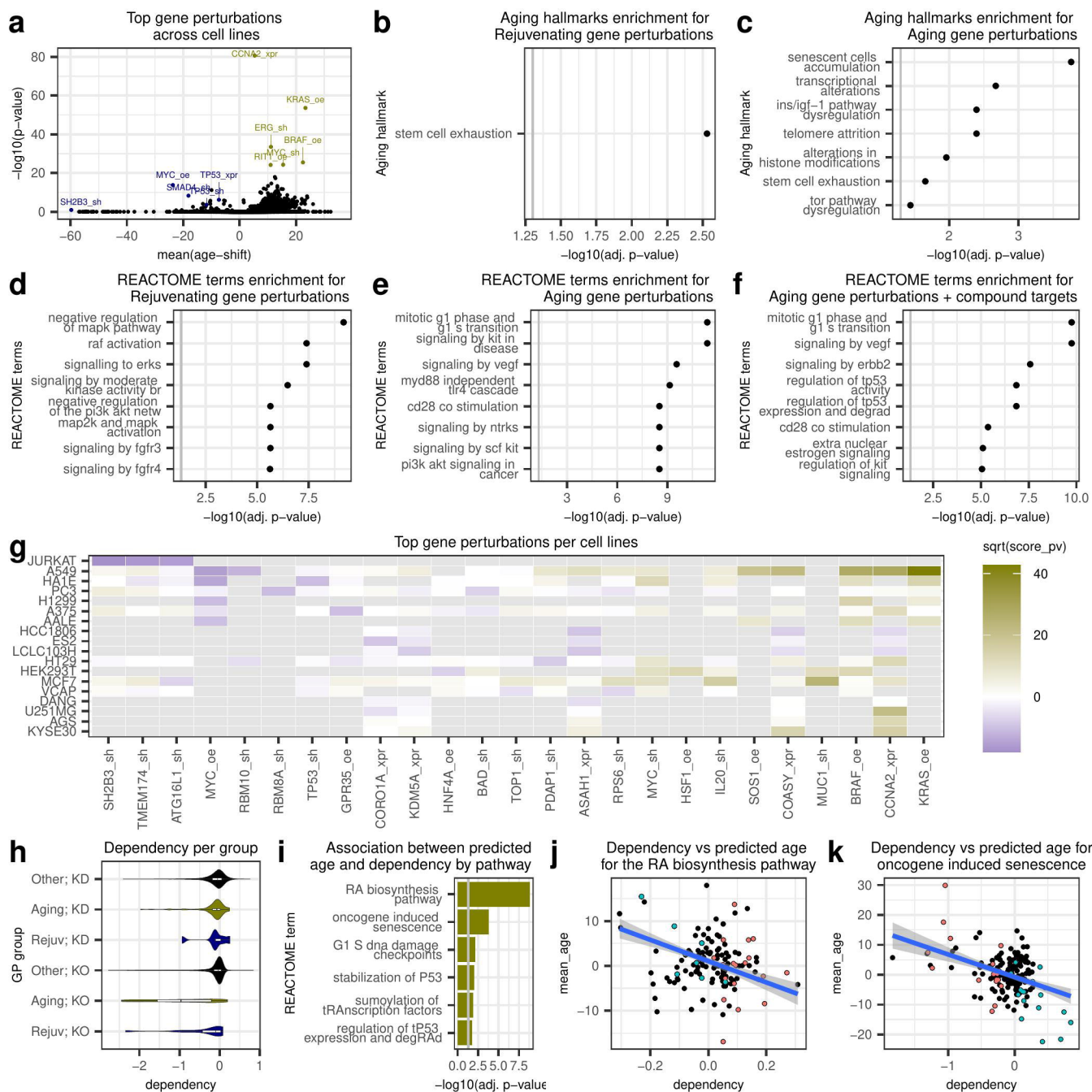

a
Top compounds across cell lines
−log10(p−value)
mean(age−shift)
mitoxantrone
PD−0325901
gemcitabine
mitomycin−c
daunorubicin
doxorubicin
repsox
b
% of senesc. induc. compounds per group
c
% of reprog. induc. compounds per group
d
% of reprog. and sen−esc.−induc. compounds
group
Rejuvenating
Other
Aging
Percent
e
MOA Aging compounds
Topoisomerase inhibitor
CDK inhibitor
MDM inhibitor
CHK inhibitor
DNA synthesis inhibitor
IMPDH inhibitor
Microtubule inhibitor
PKC activator
PARP inhibitor
RNA synthesis inhibitor
Protein synthesis inhibitor
Farnesyltransferase inhibitor
MTOR inhibitor
IKK inhibitor
Thymidylate synthase inhibitor
Akt inhibitor
Aurora kinase inhibitor
KIT inhibitor
P38 MAPK inhibitor
Ribonucleotide reductase inhibitor
Dihydrofolate reductase inhibitor
Steroidal progestin
PI3K inhibitor
Ribosomal protein inhibitor
RNA polymerase inhibitor
STAT inhibitor
−log10(adj. p−value)
f
MOA Reguvenating compounds
HDAC inhibitor
MEK inhibitor
Cardiac glycoside
Proteasome inhibitor
PI3K inhibitor
GSK−3 inhibitor
−log10(adj. p−value)
g
Top compounds per cell lines
MDAMB231
HA1E
PC3
A549
XC.L10
A375
MCF10A
HEPG2
U2OS
HME1
XC.R10
VCAP
XC.L100
MCF7
HS578T
HCT116
THP1
NPC
NCIH2110
SHSY5Y
HL60
NEU
MG−132
bortezomib
trichostatin−a
vorinostat
PD−0325901
panobinostat
PMA
AZD−9291
AZD−3514
TAS−103
camptothecin
CGP−60474
AGN−192403
pralatrexate
gemcitabine
topotecan
mitoxantrone
AMG−232
sqrt(score_pv)
h
Dependency per group
Compound group
Other
Aging
Rejuvenating
dependency
i
Association between predicted age and dependency by MOA
moa
MDM inhib.
IMPDH inhib.
Adenos. rec. antag.
SIRT inhib.
Fat. acid synt. inhib.
−log10(adj. p−value)
j
Association between predicted age and dependency by compound
drug
AMG−232
CGM−097
BRL−50481
mycophenolate−mofetil
opipramol
−log10(adj. p−value)
k
Dependency vs predicted age for MDM inhibitors
mean_age
dependency

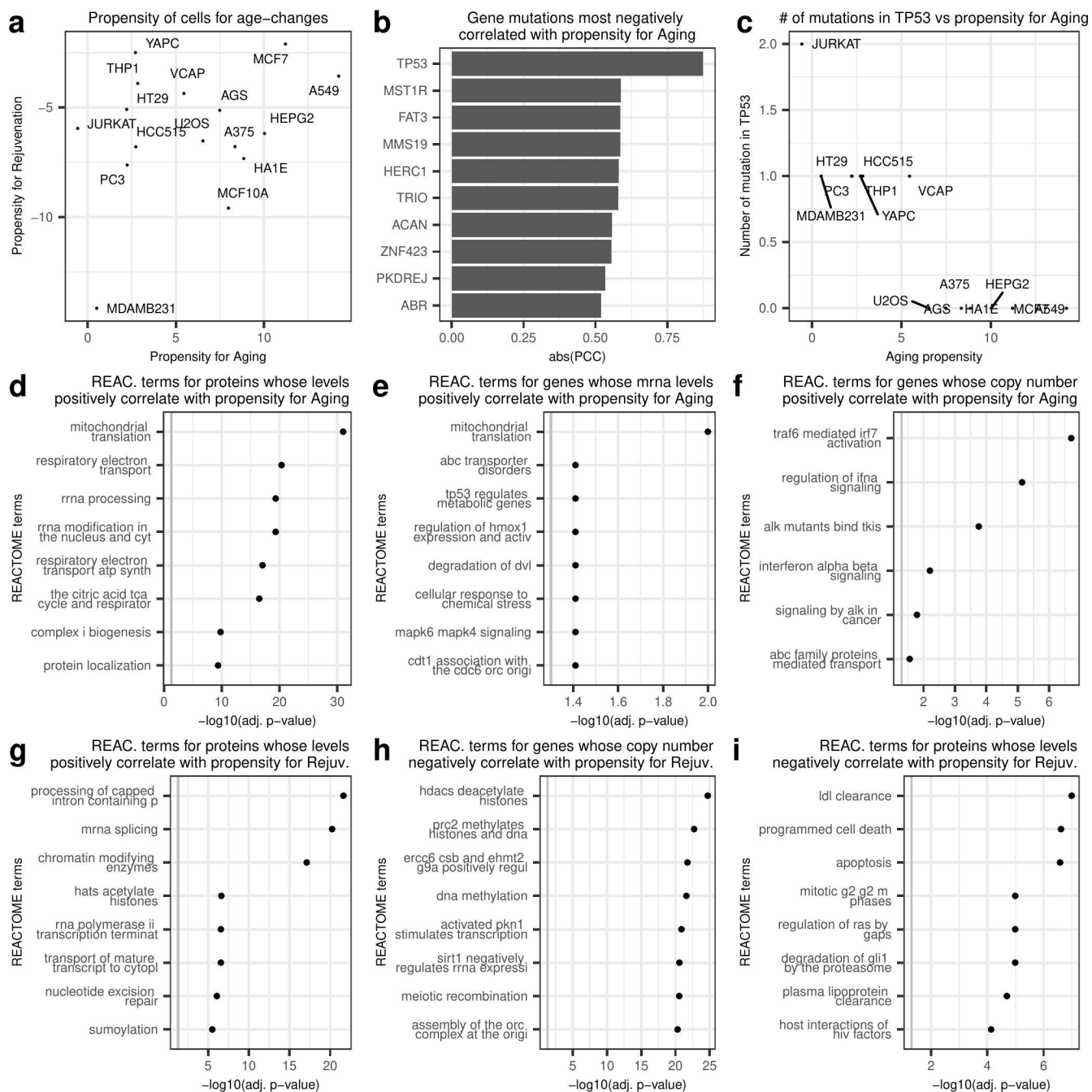
a
Anti−cancer drugs enriched in age−modulatory compounds
Cancer type
Acute Myeloid Leukemia (AML)
Ovarian Cancer
Metastatic Breast Cancer
Breast Cancer
Small Cell Lung Cancer (SCLC)
Acute Lymphocytic Leukemia (ALL)
Non−Small Cell Lung Carcinoma (NSCLC)
Ewing s Sarcoma
Non−Hodgkin s Lymphoma (NHL)
Sarcoma, Osteogenic
Head and Neck Carcinoma
Soft Tissue Sarcoma (STS)
Acute Lymphoblastic Leukaemias (ALL);
Malignant Neoplasm of Stomach
Lymphoma, Hodgkins
Advanced Renal Cell Carcinoma
Acute Lymphoblastic Leukaemias (ALL)
Multiple Myeloma (MM)
0
3
6
9
−log10(padj)


b
Top cancer−specific
senescence−inducing compounds
Compound
pralatrexate
exifone
AZD−8055
alisertib
brequinar
clofarabine
KPT−330
chaetocin
BRD−K23488072
trifluridine
0
50
100
150
200
score_pv


c
Mean age effect for cancer
and normal cell lines
Mean age effect for normal cell lines
5
0
−5
−10
clofarabine
pralatrexate
trifluridine
KPT−330
AZD−8055
brequinar
alisertib
BRD−K23488072
chaetocin
exifone
0
10
20
Mean age effect for tumor cell lines


Senescence induction in responder A375
a
Ctrl
Prltx
Doxo
b
CDKN1A mRNA (F.C)
IL6 mRNA (F.C)
.075
c
Cell proliferation (%)
cells (%)
EdU+
EdU-
d
LMNB1
CDKN1A
GAPDH
67 kDa
21 kDa
36 kDa
Senescence induction in non-responder MDA-MB-231
e
f
ns
g
h
Stemness induction in responder PC3
i
OCT4 mRNA (F.C)
SOX2 mRNA (F.C)
NANOG mRNA (F.C)
Piplng
j
OCT 3/4
TUBA1A
38 kDa
50 kDa


a
Top gene perturbations across cell lines
−log10(p−value)
mean(age−shift)
CCNA2_xpr
KRAS_oe
ERG_sh
BRAF_oe
MYC_sh
RIN1_oe
MYC_oe
TP53_xpr
SMAD4_sh
TP53_sh
SH2B3_sh
b
Aging hallmarks enrichment for Rejuvenating gene perturbations
Aging hallmark
stem cell exhaustion
−log10(adj. p−value)
c
Aging hallmarks enrichment for Aging gene perturbations
Aging hallmark
senescent cells accumulation
transcriptional alterations
ins/igf−1 pathway dysregulation
telomere attrition
alterations in histone modifications
stem cell exhaustion
tor pathway dysregulation
−log10(adj. p−value)
d
REACTOME terms enrichment for Rejuvenating gene perturbations
REACTOME terms
negative regulation of mapk pathway
raf activation
signalling to erks
signaling by moderate kinase activity br
negative regulation of the pi3k akt netw
map2k and mapk activation
signaling by fgfr3
signaling by fgfr4
−log10(adj. p−value)
e
REACTOME terms enrichment for Aging gene perturbations
REACTOME terms
mitotic g1 phase and g1 s transition
signaling by kit in disease
signaling by vegf
myd88 independent tlr4 cascade
cd28 co stimulation
signaling by ntrks
signaling by scf kit
pi3k akt signaling in cancer
−log10(adj. p−value)
f
REACTOME terms enrichment for Aging gene perturbations + compound targets
REACTOME terms
mitotic g1 phase and g1 s transition
signaling by vegf
signaling by erbb2
regulation of tp53 activity
regulation of tp53 expression and degrad
cd28 co stimulation
extra nuclear estrogen signaling
regulation of kit signaling
−log10(adj. p−value)
g
Top gene perturbations per cell lines
sqrt(score_pv)
JURKAT
A549
HA1E
PC3
H1299
A375
AALF
HCC1806
ES2
LCLC103H
H129
HEK293T
MCF7
VCAP
DANG
U251MG
AGS
KYSE30
SH2B3_sh
TMEM174_sh
ATG16L1_sh
MYC_oe
RBM10_sh
RBM8A_sh
TP53_sh
GPR35_oe
CORO1A_xpr
KDM5A_xpr
HNF4A_oe
BAD_sh
TOP1_sh
PDAP1_sh
ASAH1_xpr
RPS6_sh
MYC_sh
HSF1_oe
IL20_sh
SOS1_oe
COASY_xpr
MUC1_sh
BRAF_oe
CCNA2_xpr
KRAS_oe
h
Dependency per group
GP group
Other; KD
Aging; KD
Rejuv; KD
Other; KO
Aging; KO
Rejuv; KO
dependency
i
Association between predicted age and dependency by pathway
REACTOME term
RA biosynthesis pathway
oncogene induced senescence
G1 S dna damage checkpoints
stabilization of P53
sumoylation of tRAnscription factors
regulation of tP53 expression and degRAd
−log10(adj. p−value
j
Dependency vs predicted age for the RA biosynthesis pathway
mean_age
dependency
k
Dependency vs predicted age for oncogene induced senescence
mean_age
dependency


a
Propensity of cells for age−changes
Propensity for Rejuvenation
Propensity for Aging
YAPC
MCF7
THP1
VCAP
A549
HT29
AGS
JURKAT
U2OS
HEPG2
HCC515
A375
HA1E
PC3
MCF10A
MDAMB231
b
Gene mutations most negatively correlated with propensity for Aging
TP53
MST1R
FAT3
MMS19
HERC1
TRIO
ACAN
ZNF423
PKDREJ
ABR
abs(PCC)
c
# of mutations in TP53 vs propensity for Aging
Number of mutation in TP53
Aging propensity
JURKAT
HT29
HCC515
PC3
THP1
VCAP
MDAMB231
YAPC
A375
HEPG2
U2OS
AGS
HA1E
MCF7
A549
d
REAC. terms for proteins whose levels positively correlate with propensity for Aging
REACTOME terms
mitochondrial translation
respiratory electron transport
rrna processing
rrna modification in the nucleus and cyt
respiratory electron transport atp synth
the citric acid tca cycle and respirator
complex i biogenesis
protein localization
−log10(adj. p−value)
e
REAC. terms for genes whose mrna levels positively correlate with propensity for Aging
REACTOME terms
mitochondrial translation
abc transporter disorders
tp53 regulates metabolic genes
regulation of hmox1 expression and activ
degradation of dvl
cellular response to chemical stress
mapk6 mapk4 signaling
cdt1 association with the cdc6 orc origi
−log10(adj. p−value)
f
REAC. terms for genes whose copy number positively correlate with propensity for Aging
REACTOME terms
traf6 mediated irf7 activation
regulation of ifna signaling
alk mutants bind tkis
interferon alpha beta signaling
signaling by alk in cancer
abc family proteins mediated transport
−log10(adj. p−value)
g
REAC. terms for proteins whose levels positively correlate with propensity for Rejuv.
REACTOME terms
processing of capped intron containing p
mrna splicing
chromatin modifying enzymes
hats acetylate histones
rna polymerase ii transcription terminat
transport of mature transcript to cytopl
nucleotide excision repair
sumoylation
−log10(adj. p−value)
h
REAC. terms for genes whose copy number negatively correlate with propensity for Rejuv.
REACTOME terms
hdacs deacetylate histones
prc2 methylates histones and dna
ercc6 csb and ehmt2 g9a positively regul
dna methylation
activated pkn1 stimulates transcription
sirt1 negatively regulates rrna expressi
meiotic recombination
assembly of the orc complex at the origi
−log10(adj. p−value)
i
REAC. terms for proteins whose levels negatively correlate with propensity for Rejuv.
REACTOME terms
ldl clearance
programmed cell death
apoptosis
mitotic g2 g2 m phases
regulation of ras by gaps
degradation of gli1 by the proteasome
plasma lipoprotein clearance
host interactions of hiv factors
−log10(adj. p−value)