

Leveraging metabolic modeling and machine learning to uncover modulators of quiescence depth

Alec Eames ^a and Sriram Chandrasekaran ^{a,b,c,d,*}

^aDepartment of Biomedical Engineering, University of Michigan, Ann Arbor, MI 48109, USA

^bProgram in Chemical Biology, University of Michigan, Ann Arbor, MI 48109, USA

^cRogel Cancer Center, University of Michigan Medical School, Ann Arbor, MI 48109, USA

^dCenter for Bioinformatics and Computational Medicine, University of Michigan, Ann Arbor, MI 48109, USA

*To whom correspondence should be addressed: Email: csriram@umich.edu

Edited By: Christian Metallo

Abstract

Quiescence, a temporary withdrawal from the cell cycle, plays a key role in tissue homeostasis and regeneration. Quiescence is increasingly viewed as a continuum between shallow and deep quiescence, reflecting different potentials to proliferate. The depth of quiescence is altered in a range of diseases and during aging. Here, we leveraged genome-scale metabolic modeling (GEM) to define the metabolic and epigenetic changes that take place with quiescence deepening. We discovered contrasting changes in lipid catabolism and anabolism and diverging trends in histone methylation and acetylation. We then built a multi-cell type machine learning model that accurately predicts quiescence depth in diverse biological contexts. Using both machine learning and genome-scale flux simulations, we performed high-throughput screening of chemical and genetic modulators of quiescence and identified novel small molecule and genetic modulators with relevance to cancer and aging.

Keywords: quiescence deepening, genome-scale models, machine learning, cancer, aging

Significance Statement

Quiescence is a sleep-like cell state important for tissue regeneration. Quiescence exists on a spectrum, from shallow to deep quiescence, based on a cell's ability to proliferate. Through computational modeling, we simulated metabolic changes associated with quiescence deepening. We also developed a machine learning model that accurately predicts quiescence depth in multiple cell types and quiescence induction methods. Our models predicted genes and small molecules capable of tuning the depth of quiescence. These modulating factors show a striking association with cancer initiation and aging.

Introduction

Cellular quiescence is a state of reversible dormancy in which cell division is halted but able to resume if needed. As a common somatic cell state, appropriate transition into and out of quiescence is essential for proper tissue homeostasis (1, 2). Quiescence is induced in stem cells, for instance, where it shields cells from environmental stress. In response to tissue injury, however, activation of quiescent stem cells is key to regenerating damaged tissue (3, 4).

Rather than a uniform state, quiescence is increasingly viewed as a continuum between shallow and deep quiescence, akin to light and deep sleep, requiring different potencies of growth stimulation to induce proliferation. Cells in shallower quiescence respond more easily to growth signals, whereas cells in deeper quiescence are more difficult to activate (5). This phenomenon is observed in a wide range of contexts. Cells cultured under serum starvation or contact inhibition become progressively

desensitized to growth signals (6). Likewise, quiescent liver cells display impaired regenerative capacity with age, reflective of a deeper quiescence (7).

Tuning the depth of quiescence is critical in a range of physiological contexts. Muscle stem cells, for instance, shift to a shallower state of quiescence in response to tissue injury or stress (8), and certain neural stem cells (NSCs) exit quiescence upon nutritional stimulus (9). Additionally, quiescence depth is dysregulated in a range of diseases and during aging. In cancer, malignant cells can become quiescent, where they evade most cancer treatments (10), and with age, quiescence deepens across various tissues and cell types, impairing tissue regeneration and compromising cellular resilience (5). In light of this, better understanding the factors modulating quiescence would shed light on an important facet of quiescent cell biology and may have therapeutic relevance to the variety of disease states in which quiescence depth is dysregulated. These factors, however, remain

Competing Interest: The authors declare no competing interest.

Received: July 1, 2023. **Accepted:** December 28, 2023

© The Author(s) 2024. Published by Oxford University Press on behalf of National Academy of Sciences. This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs licence (<https://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits non-commercial reproduction and distribution of the work, in any medium, provided the original work is not altered or transformed in any way, and that the work is properly cited. For commercial re-use, please contact journals.permissions@oup.com

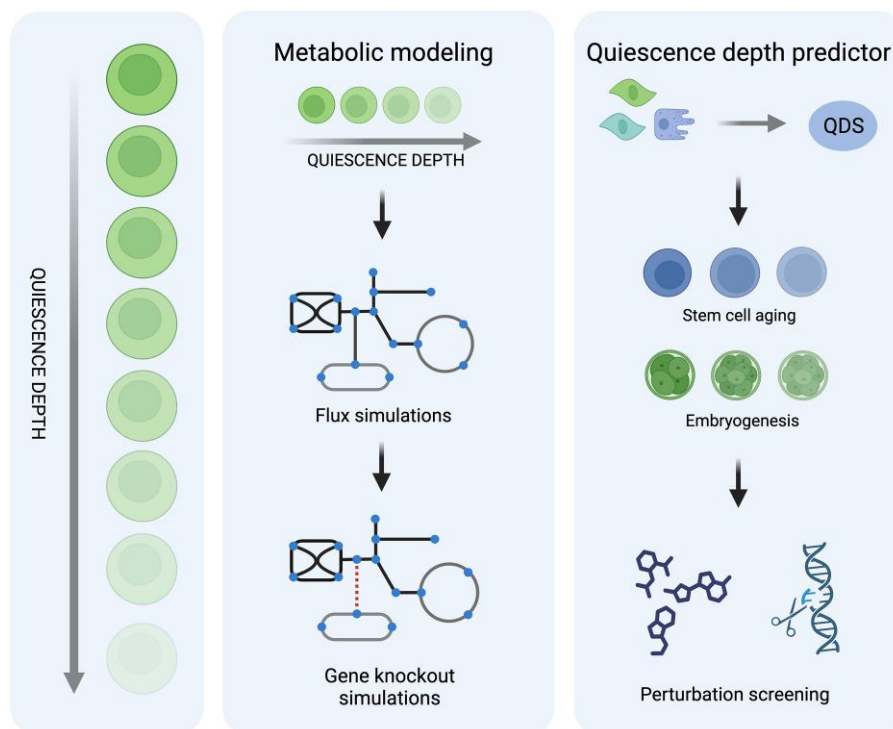


Fig. 1. Graphical overview. In this article, we take a multifaceted approach to identifying modulators of quiescence depth: (i) metabolic modeling and flux simulations that enable in silico genetic knockouts and (ii) a multi-cell type transcriptomics quiescence depth predictor that facilitates screening chemical and genetic perturbations for their effect on quiescence depth. In doing so, we simulate metabolic and epigenetic features of quiescence deepening and explore quiescence depth in embryogenesis and aging. Predictions are validated using a cancer genomics database.

poorly defined. In this study, we take a multifaceted approach to uncover modulators of quiescence depth encompassing genome-scale metabolic model (GEM) and machine learning predictors of quiescence.

Genome-scale metabolic models are mathematical representations of metabolic networks encompassing an organism's entire genome (11). By incorporating constraints, including physicochemical and environmental limitations, GEMs can simulate fluxes through metabolic reactions and accurately predict responses to genetic perturbations (12). GEMs can also incorporate omics data to further constrain the set of allowable fluxes (13). For instance, Frezza et al. (14) created a tissue-specific metabolic model using transcriptomics data, and then accurately predicted how metabolism was rewired in the absence of a key citric acid cycle enzyme. GEMs have also been used to define the metabolic signature of quiescent cancer cells through omics data incorporation (15).

Here, we leveraged metabolic models to simulate the metabolic changes underlying quiescence deepening and perform in silico gene knockouts to identify genes capable of shifting quiescence depth toward a shallower or deeper state. In doing so, we highlighted metabolic and epigenetic trends associated with quiescence deepening and built a metabolic flux-based quiescence depth predictor that accommodates multiple types of input data. Metabolism and epigenetics are profoundly important to cellular biology, yet the metabolic and epigenetic trends associated with quiescence deepening remain unexplored, so this analysis provides a window into critical yet undefined aspects of quiescence.

In addition, we developed a novel transcriptomics-based quiescence depth predictor built using multiple quiescence induction methods and multiple cell types and validated its accuracy in diverse contexts. Using this predictor, we performed high-throughput screening of chemical and genetic

perturbations to identify interventions that robustly altered quiescence depth. We also explored quiescence depth in embryogenesis and stem cell aging at the single-cell level. We validated our predictions using cancer genomics data (Fig. 1).

Results

Metabolic modeling simulates metabolic fluxes during quiescence deepening

First, we simulated metabolic fluxes for progressively deeper stages of quiescence by incorporating transcriptomics from nine stages of quiescence depth into a GEM containing 2,953 genes and 8,377 metabolites organized into 13,028 reactions (16). Quiescence deepening was induced in rat fibroblasts through serum starvation, and quiescence depth was quantified as the days since serum starvation began, denoted as quiescence depth score (QDS) (5). Each quiescence depth score had three replicates for a total of 27 samples. To incorporate transcriptomics, we used Integrative Network Inference for Tissues (INIT), a well-established tool for creating transcriptomics-constrained metabolic models (17). Then we randomly sampled from the flux solution space (Fig. 2A). Similar results were obtained using linear Integrative Metabolic Analysis Tool (iMAT) (18) and parsimonious flux balance analysis (see Methods).

We then identified which metabolic reactions corresponded most strongly to quiescence deepening by correlating reaction fluxes with quiescence depth and highlighting the reactions with the strongest association. The fluxes of these reactions show a strong correspondence with quiescence depth (Fig. 2E and F). To better understand the metabolic pathways represented by these reactions, each reaction's association with a subsystem

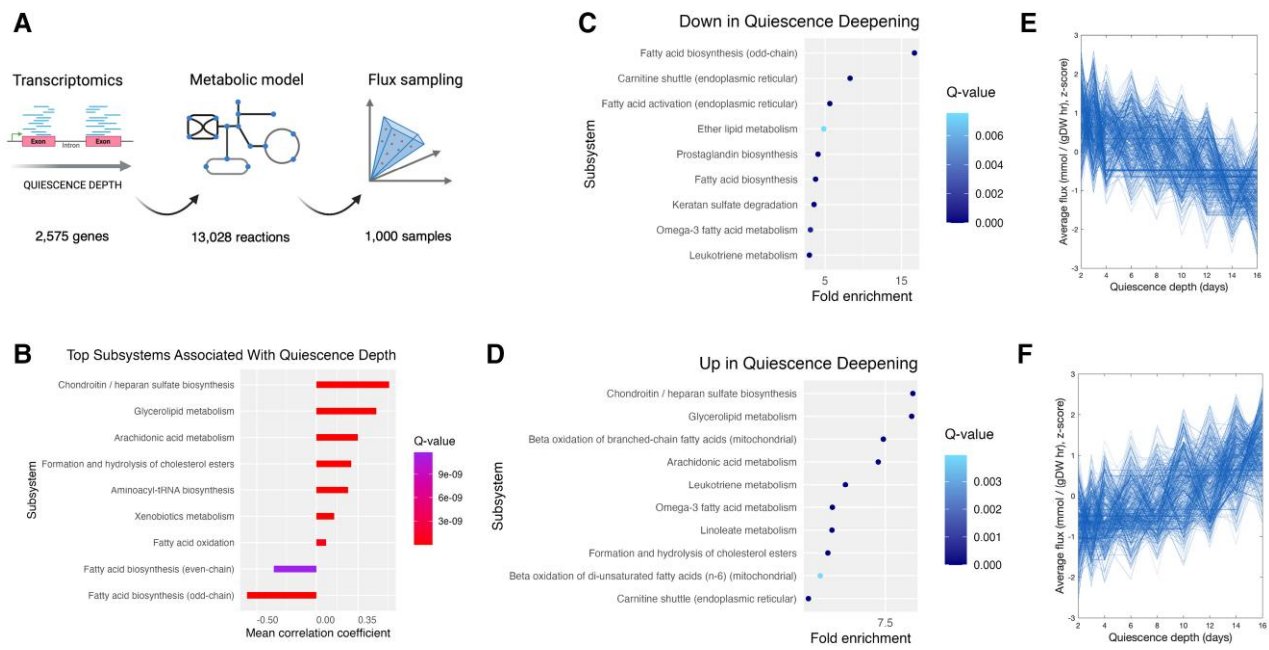


Fig. 2. Metabolic pathway analysis of quiescence deepening. A) Workflow for simulating metabolic reaction fluxes. Transcriptomics from nine stages of quiescence depth were integrated into a metabolic model (Rat-GEM) to constrain the set of allowable fluxes (using INIT). Then, we used CHRR to randomly sample from the set of potential fluxes for each quiescence depth. B) Top subsystems associated with quiescence depth. For each subsystem, each reaction in the subsystem was correlated with quiescence depth and the average Pearson correlation was computed. Subsystems with the highest average Pearson correlation coefficients are shown here (Q-value < 1E-5, one-sample t test). A positive mean correlation coefficient implies a positive association with quiescence depth while a negative mean correlation coefficient indicates a negative association with quiescence depth. C) Subsystem enrichment analysis of reactions negatively correlating with QDS at a P-value of < 0.15. Subsystems (pathways) ranked by fold enrichment are shown here along with an enrichment P-value calculated using the hypergeometric function and corrected with the Benjamini–Hochberg adjustment (Q-value < 0.01 subsystems shown here). D) Subsystem enrichment analysis of reactions positively correlating with QDS at a P-value of < 0.15. Subsystems (pathways) ranked by fold enrichment are shown here along with an enrichment P-value calculated using the hypergeometric function and corrected with the Benjamini–Hochberg adjustment (Q-value < 0.01 subsystems shown here). Top 10 subsystems meeting criteria are shown here, ranked by fold enrichment. E) Normalized reaction fluxes across quiescence depths for all reactions negatively correlating with QDS at a P-value of < 0.15. F) Normalized reaction fluxes across quiescence depths for all reactions positively correlating with QDS at a P-value of < 0.15.

in the metabolic model was leveraged. Using this information, we performed an enrichment analysis to identify which metabolic subsystems were overrepresented in the reactions positively correlating to quiescence depth (increasing with quiescence depth) and the reactions negatively correlating to quiescence depth (decreasing with quiescence depth) (Fig. 2C and D). The basis for this fold enrichment is the frequency of a given subsystem in the differential fluxes compared to the frequency of a given subsystem in all reactions in the metabolic model. Glycosaminoglycan, arachidonic acid, and linoleate metabolism were a few of the subsystems found to rise with quiescence deepening, whereas fatty acid biosynthesis, prostaglandin biosynthesis, and omega-3 metabolism were predicted to fall.

A further pathway analysis was then performed in which each metabolic subsystem's average correlation with quiescence depth was calculated, and subsystems with the highest absolute correlation were identified (Fig. 2B). This revealed many of the same trends, with lipid metabolism subsystems conspicuous.

Given the enrichment of lipid metabolism pathways in these subsystems, we decided to delve deeper into these pathways specifically. To do so, we examined a comprehensive set of lipid-related metabolic subsystems and corroborated the trends using transcriptomics data.

Lipid anabolism and catabolism display largely opposing trends during quiescence deepening

The most striking metabolic pattern observed during quiescence deepening was the diverging trends of lipid anabolism and

catabolism. As quiescence deepens, beta-oxidation fluxes substantially increase while fatty acid and cholesterol biosynthesis fluxes decline, albeit in a less uniform fashion (Fig. 3A and B), suggesting a metabolic shift from fatty acid biosynthesis to fatty acid oxidation takes place during quiescence deepening. We highlight representative metabolic subsystems and reactions in Fig. 3C and D for lipid anabolism and catabolism, respectively, using Escher plots (19). The full-size metabolic network plots are shown in Figs. S2 and S3.

This metabolic shift aligns strongly with experimental metabolomics data from Yao et al. (20) in which quiescent cells displayed higher rates of carnitine-shuttle-dependent fatty acid oxidation and lower rates of lipid biosynthesis compared to proliferating controls. It also aligns with transcriptomics data used for model reduction. To corroborate the trends with transcriptomics data, we identified genes related to lipid biosynthesis and lipid catabolism, and then computed their correlation with quiescence depth (Fig. 3E and F). For lipid biosynthesis genes, we see a modest majority decline with quiescence deepening whereas for lipid catabolism, a more substantial majority increases with quiescence depth, both aligning with the trends seen in the fluxes.

Epigenetic simulations reveal diverging trends in histone acetylation and methylation as quiescence deepens

Epigenetic modifications play a profound role in regulating a cell's gene expression, and their nature is highly dependent

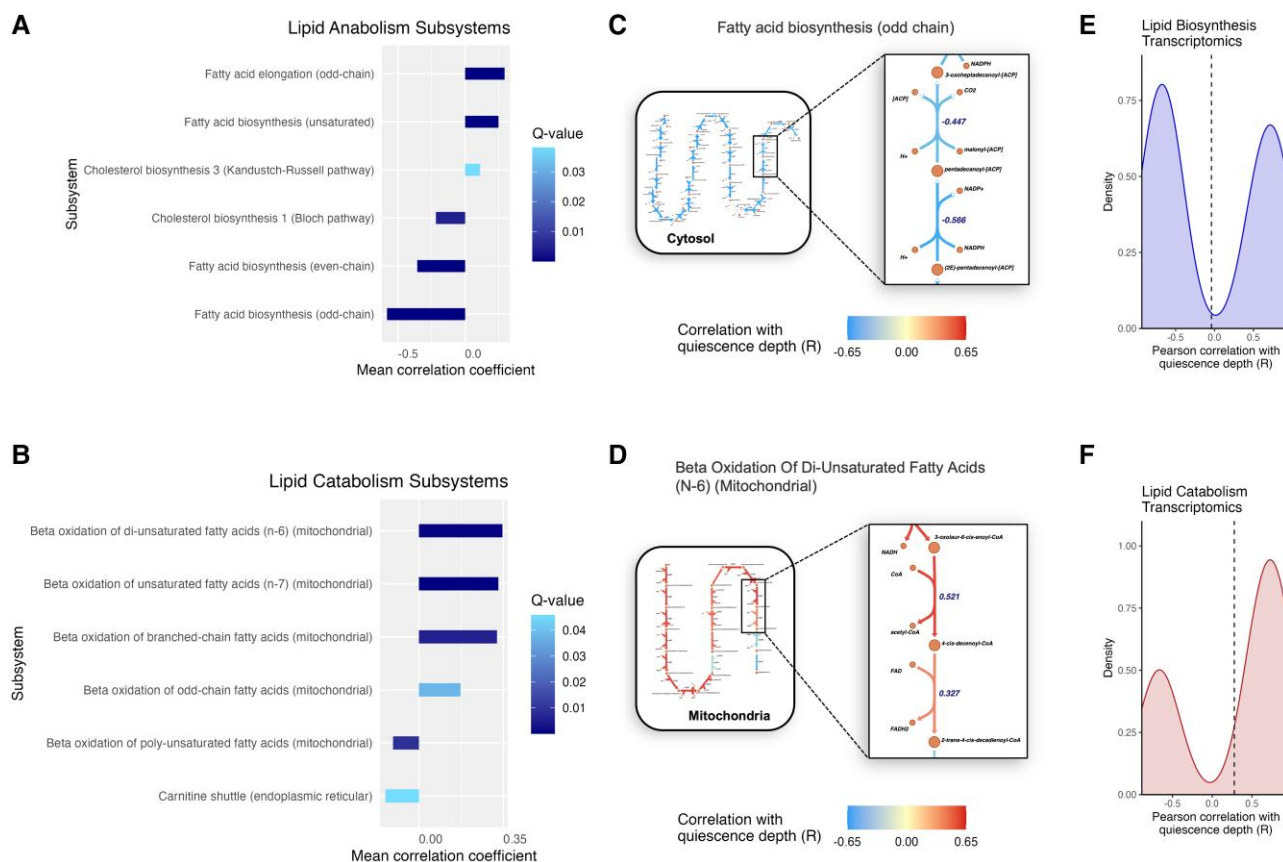


Fig. 3. Lipid anabolism and lipid catabolism display largely opposing trends during quiescence deepening. A) Mean Pearson correlation coefficient of all reactions with QDS for lipid anabolism subsystems. Only subsystems with a mean correlation coefficient different from zero (Q-value < 0.05, one-sided t test) are shown here. B) Mean Pearson correlation coefficient of all reactions with QDS for lipid catabolism subsystems. Only subsystems with a mean correlation coefficient different from zero (Q-value < 0.05, one-sided t test) are shown here. C) Metabolic pathway plot for example lipid anabolism subsystem: fatty acid biosynthesis (odd-chain). Reactions are colored by the Pearson correlation coefficient with quiescence depth. D) Metabolic pathway plot for example lipid catabolism subsystem: beta oxidation of di-unsaturated fatty acids (N-6) (mitochondrial). Reactions are colored by the Pearson correlation coefficient with quiescence depth. E) Distribution of Pearson correlation coefficients with quiescence depth for transcripts belonging to the GO biological process “lipid biosynthesis.” Only correlation coefficients with Q-value < 0.05 are shown here. F) Distribution of Pearson correlation coefficients with quiescence depth for transcripts belonging to the GO biological process “lipid catabolism.” Only correlation coefficients with Q-value < 0.05 are shown here.

on the metabolic state of the cell. The metabolite acetyl-CoA, for instance, serves as the substrate for histone acetylation, a histone modification that activates gene transcription, while the metabolite S-adenosylmethionine (SAM) serves as the carbon donor in histone methylation (21, 22) (Fig. 4A). Given the interplay between metabolism and epigenetics, metabolic models enable the simulation of histone modifications based on the abundance and flux of metabolic precursors (23–25).

Leveraging this ability, we simulated histone acetylation and methylation as quiescence transitions from a shallow to a deep state. Remarkably, we observed diverging trends in these two histone modifications: histone acetylation is predicted to fall during quiescence deepening (Fig. 4B) while histone methylation is predicted to rise (Fig. 4D). In agreement with these results, histone acetylation is known to fall with quiescence induction (26).

What metabolic pathways could be driving these trends? To answer this question, we knocked out each subsystem in the metabolic model and observed the resulting epigenetic trend. For histone acetylation, the downward trend was abolished most strongly by knocking out fatty acid activation and carnitine-shuttle-related reactions—pathways known to impact histone acetylation in other contexts (27, 28) (Fig. 4C). Given the steady increase in carnitine-shuttle-dependent fatty acid

oxidation predicted during quiescence deepening, this suggests that increased fatty acid oxidation may be responsible for mediating the downward trend in histone acetylation. For histone methylation, knockdown of peptide metabolism and several lipid-related subsystems—both anabolic and catabolic—had the strongest impact in abolishing the upward trend (Fig. 4E), implying a link may exist between rising histone methylation and the predicted shift from fatty acid anabolism to catabolism during quiescence deepening.

In silico gene knockout simulations predict a range of quiescence depth-altering genes that align with experimental data

Metabolic models enable the large-scale simulation of genetic perturbations by applying known information about an organism’s metabolic networks and their relationship to genes (13). Leveraging this functionality, we performed a gene knockout screening to identify genes that may modulate quiescence depth.

To accomplish this, we employed the algorithm robust metabolic transformation algorithm (rMTA), which identifies genes whose knockout shifts the metabolic state of a cell from a source state to a target state (29). We simulated metabolic fluxes for shallow

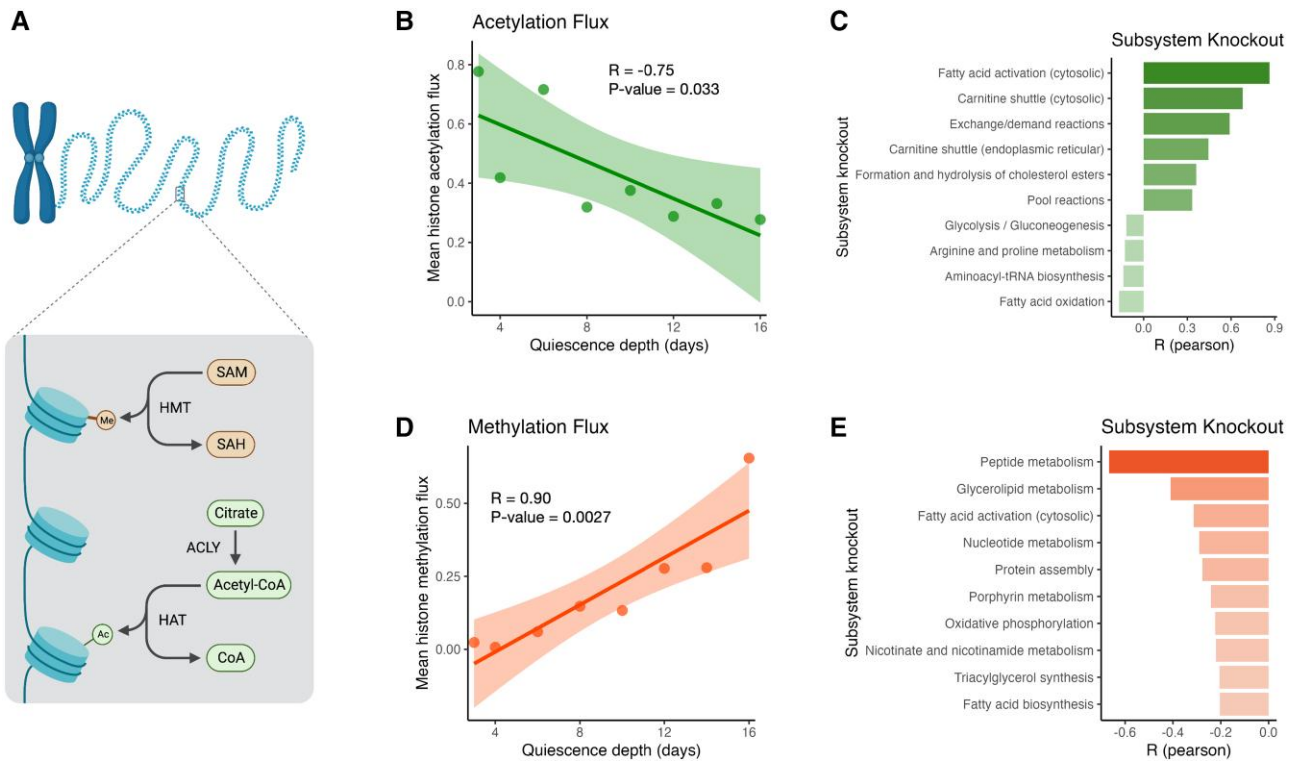


Fig. 4. Quiescence deepening is associated with diverging trends in epigenetic histone acetylation and methylation. A) Diagram of chromatin highlighting histone post-translational modifications and their relationship to metabolism. SAM is the substrate for histone methyltransferases (HMT) to methylate histone residues and convert SAM to S-adenosylhomocysteine (SAH). Acetyl-CoA is the substrate for histone acetyltransferases (HAT) to acetylate histone residues with CoA as a byproduct. Citrate is converted to acetyl-CoA by ATP citrate lyase (ACLY). These reactions are simulated in the metabolic model. B) Simulated mean nuclear histone acetylation flux vs. quiescence depth reveals a downward trend in histone acetylation as quiescence deepens. Pearson correlation performed between quiescence depth and mean nuclear histone acetylation flux. C) Top subsystems whose knockout reverse the acetylation trend, ranked by Pearson correlation between QDS and histone acetylation upon subsystem knockout. D) Simulated mean nuclear histone methylation flux vs. quiescence depth reveals an upward trend in histone methylation with increasing quiescence depth. Pearson correlation performed between quiescence depth and mean nuclear histone methylation flux. E) Top subsystems whose knockout reverses the methylation trend, ranked by Pearson correlation between QDS and histone methylation upon subsystem knockout.

quiescence and deep quiescence through flux sampling, then performed two screenings: one for genes whose knockout shifts the cell from shallow to deep quiescence and another for genes whose knockout shifts the cell from deep to shallow quiescence (Fig. 5A).

To validate the results of these predictions, we compared the predictions with known oncogenes and tumor suppressors from the CancerMine database (30). This comparison relied on the following principle: genes whose knockdown lowered quiescence depth (“quiescence deepeners”) were considered validated if they were annotated in the database as tumor suppressors (genes whose downregulation promotes cancer growth) and genes whose knockdown deepened quiescence depth (“quiescence reducers”) were considered validated if they were annotated as oncogenes or cancer drivers (genes whose downregulation lowers cancer growth). This validation approach was used because of the sparsity of studies screening genes directly for quiescence depth and proliferation’s strong association with quiescence depth—shallower quiescent cells are more easily able to proliferate and several genes that alter quiescence depth also affect proliferation in nonquiescent cells, such as genes in the Rb-E2F pathway (31, 32).

For genes whose knockdown was predicted to shift the cell to deeper quiescence (“quiescence reducers”), a remarkable percentage agree with experimental data. Virtually, all these genes with CancerMine annotations are considered cancer oncogenes or drivers (19/20 of the top 20 predictions with annotations, 23/25 of the top 25 predictions with annotations), which aligns

with a role in driving quiescence to a shallower state with more proliferative potential (Fig. 5B). The enrichment for cancer driver genes in these predictions out of all genes tested is shown in Fig. 5C along with the P-value calculated using the hypergeometric function. Examining genes in this category reveals that many are involved in lipid metabolism and cancer progression. The genes fatty acid synthase (FASN) and fatty acid elongase 1 (ELOVL1), for instance, are both involved in fatty acid biosynthesis and cancer initiation (33, 34), and the genes cholesterol delta-isomerase (EBP), squalene epoxidase (SQLE), and 24-dehydrocholesterol reductase (DHCR24) represent a diverse set of cholesterol biosynthesis-related oncogenes (35).

For genes whose knockdown is predicted to shift the cell toward shallower quiescence (“quiescence deepeners”), 8 of the top 21 predictions with annotations are classified as tumor suppressors. While this enrichment does not reach statistical significance out of all tested genes, one prediction is especially noteworthy for its alignment with experimental data: mitochondrial pyruvate carrier 1 (MPC1), which is listed among the top 20 genes whose knockout reduces quiescence mostly strongly out of 2,953 tested genes. This gene has direct experimental evidence for its inhibition lowering quiescence depth. Both pharmacological and genetic inhibition of MPC1 activates quiescent NSCs by shifting metabolism away from oxidative phosphorylation and toward glycolysis (36), in alignment with the simulated results. Exploring other predictions in this category shows that several

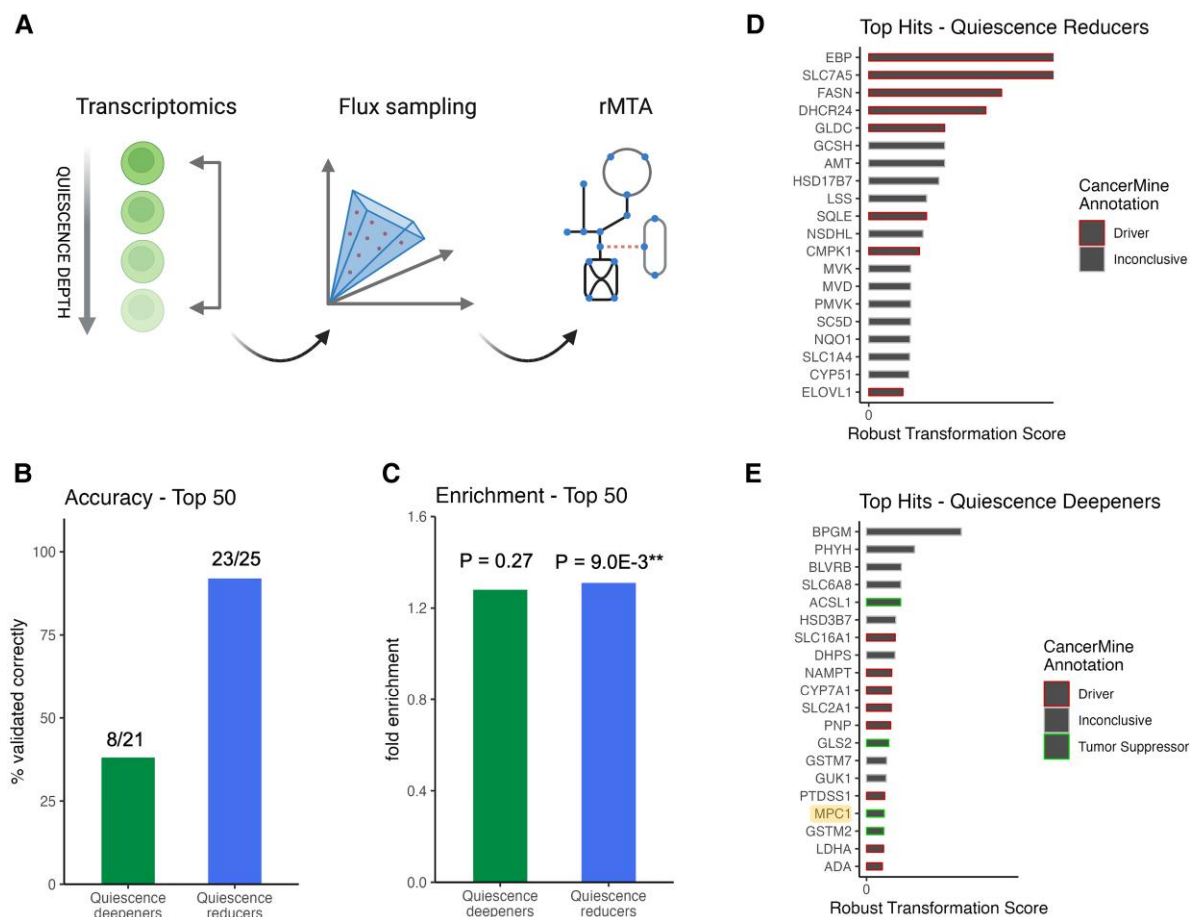


Fig. 5. Gene knockout simulations predict a range of quiescence depth modulators in alignment with experimental data. A) Overview of gene knockout simulations. Transcriptomics from shallow and deep quiescence states were used to constrain a metabolic model and perform flux simulations, enabling the simulation of the metabolic state of a cell in shallow and deep quiescence. Then, we used the rMTA to predict gene knockouts that could move a cell to (i) a shallower quiescence state or (ii) a deeper quiescence state. B) Accuracy of gene perturbation simulations for top 50 hits. To validate these results experimentally, we used a cancer genomic database called CancerMine. The CancerMine database validation relied on the following principle: genes whose knockdown lowered quiescence depth was considered validated if they were annotated as tumor suppressors (genes whose downregulation promotes cancer growth), genes whose knockdown raised quiescence were considered validated if they were annotated as oncogenes (genes whose downregulation lowers cancer growth). C) Enrichment analysis. For quiescence deepeners, we calculated the fold enrichment and P-value of tumor suppressors among the top 50 hits compared to the background list of all genes knocked out with rMTA. For quiescence reducers, we calculated the fold enrichment of cancer oncogenes or drivers compared to the background set of genes. P-values are computed using the hypergeometric function and listed above the bar chart bars. **P-value < 0.01. D) Top 20 hits of the gene perturbation simulation for quiescence reducers: The top 20 genes whose knockout is predicted to move the cells toward a deeper quiescence state are shown here (ranked by rMTA robust transformations core). Here, the CancerMine annotation described in B) is shown for each prediction. E) Top 20 hits of the gene perturbation simulation for quiescence deepeners: The top 20 genes whose knockout is predicted to move the cells toward a shallower quiescence state are shown here (ranked by rMTA robust transformations core). As before, the CancerMine annotation described in B) is shown here for each prediction. The gene MPC1 is highlighted. This gene has experimental evidence for its inhibition lowering quiescence depth (as predicted here).

genes are involved in bile acid synthesis, including 3 beta-hydroxysteroid dehydrogenase type 7 (HSD3B7) and cholesterol 7 α -hydroxylase (CYP7A1) (37), and several are involved in nucleotide metabolism such as purine nucleoside phosphorylase (PNP), guanylate kinase 1 (GUK1), and adenosine deaminase (ADA) (38–40).

The top 20 hits for both quiescence reducers and drivers along with their CancerMine annotation are shown in Fig. 5D and E. The top 50 predictions for both quiescence reducers and deepeners are listed in Table S1.

Flux-based quiescence depth predictor captures metabolic trends associated with quiescence deepening for multiple types of input data

To synthesize the metabolic variation underlying quiescence deepening into a single score, we built a machine-learning model to

predict QDS using reaction fluxes. This model captures the metabolic trends behind quiescence deepening in an easily interpretable manner. It also accommodates multiple types of input data. An elastic net architecture was chosen to build the most parsimonious model and strike a balance between underfitting and overfitting the reaction coefficients (i.e. too many and too few model coefficients).

The model's accuracy in predicting quiescence depth was validated both internally and externally. Based on internal *k*-fold cross-validation, the model accurately predicted quiescence depth across the entire range of values when trained on 90% of the data and tested on the remaining 10% (Fig. 6C). To validate the accuracy of the model externally in independent datasets, we examined six datasets involving nine distinct experimental conditions for proliferating vs. quiescent cells (42–45). Given that shallower quiescent cells are closer to proliferation than deeper quiescent cells, we reasoned that quiescent cells should have a

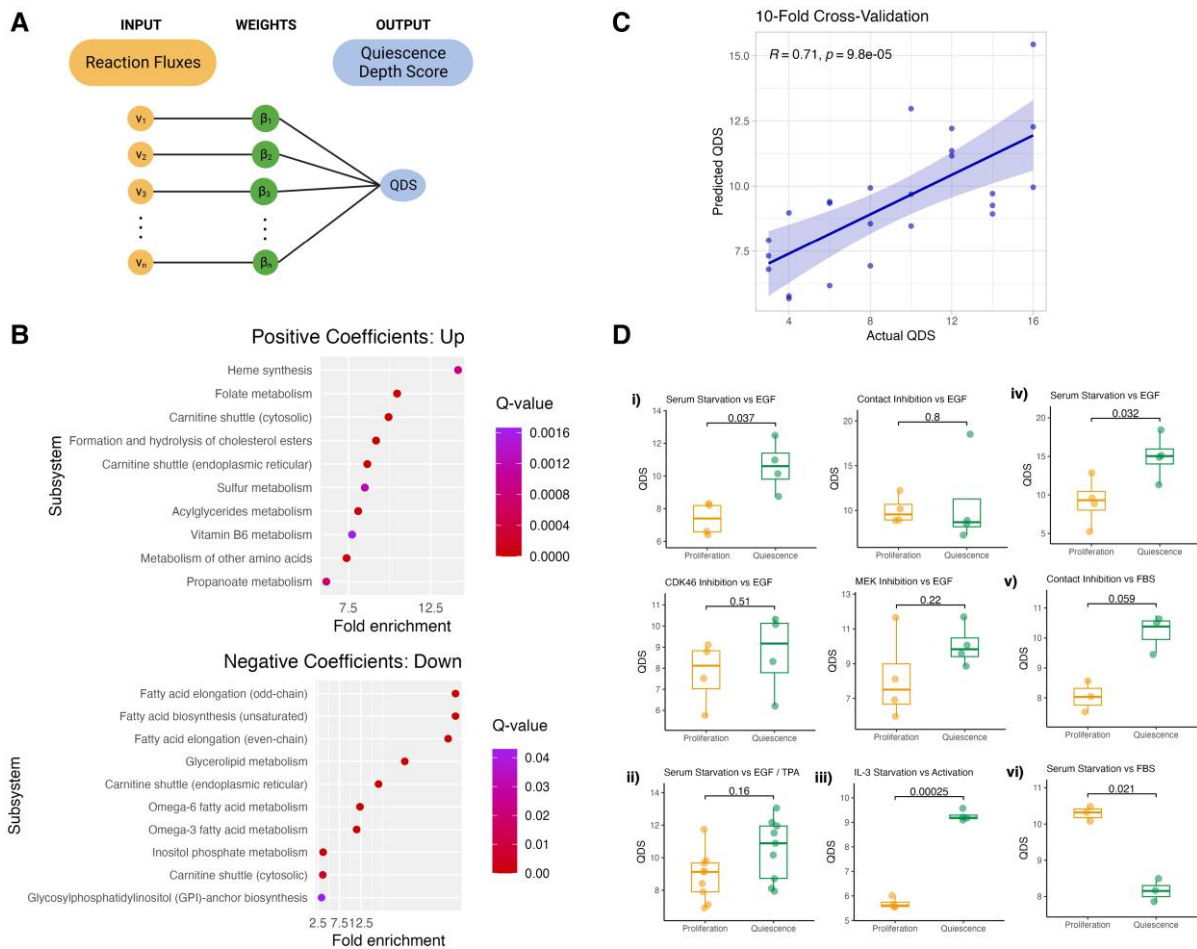


Fig. 6. Flux-based machine learning model captures metabolic signatures of quiescence deepening for multiple types of input data. A) Diagram of the machine learning model. Reaction fluxes are inputted, and the model is trained to predict QDS (days since serum starvation) using elastic net regression. B) Top subsystems enriched in the model's positive coefficients (up with quiescence deepening) and negative coefficients (down with quiescence deepening). The top 10 subsystems with $Q\text{-value} < 0.05$ for the hypergeometric fold enrichment test are shown here. C) K-fold cross-validation accuracy plot. The model is trained and optimized on 90% of the data and tested in the remaining 10% for 10 random subsets of the data. Note: all parameter are un-optimized in training set except for fold-change threshold, which is set close to one to maximize the impact of genes on fluxes. D) External validation shows that the model predicts deeper quiescence (higher QDS) for cells in a quiescent phase vs. a proliferative phase for multiple types of quiescence and proliferation induction and multiple types of input omics data (transcriptomics, translomics, and proteomics). Fluxes are generated using genes upregulated in quiescence relative to proliferation vs. fluxes generated using genes upregulated in proliferation relative to quiescence with a fold change threshold of 1.00. (i) Proliferation is induced using epidermal growth factor (EGF), horse serum, insulin, and hydrocortisone. Quiescence is induced with serum starvation, Mek inhibition, CDK46 inhibition, and contact inhibition. Transcriptomics measured. Comparisons were performed with a two-sample t test. Source: GSE122927 (41). (ii) Proliferation is induced using EGF or tissue plasminogen activator (TPA). Quiescence induced with serum starvation. Transcriptomics measured. Comparisons were performed with a two-sample t test. Source: Ref. (42). (iii) QDS for proliferating vs. quiescent cells. Proliferation is induced by IL-3. Quiescence is induced with IL-3 starvation. Proteomics measured. Comparisons were performed with a paired t test. Source: Ref. (43). (iv) QDS for proliferating vs. quiescent cells. Proliferation is induced using EGF. Quiescence induced with serum starvation. Translatomics measured. Comparisons were performed with a paired t test. Source: GSE112295 (44). (v) Proliferation is induced using fetal bovine serum (FBS). Quiescence induced with 7-day contact-inhibition. Transcriptomics measured. Comparisons were performed with a two-sample t test. Source: GSE117444 (45). (vi) Proliferation is induced using FBS. Quiescence induced with serum starvation. Transcriptomics measured. Comparisons were performed with a two-sample t test. Source: GSE74620 (46).

higher QDS relative to proliferating cells. Here, we observed a trend toward higher quiescence depth in seven of the nine datasets and a higher quiescence depth in four of the nine datasets at a P -value threshold of 0.1 (Fig. 6D) despite vastly different experimental conditions than the data the model was trained on, including interleukin 3 (IL-3) starvation in murine lymphocytes (Fig. 6D, iii) and serum starvation in human cancer cells (Fig. 6D, iv). The model, however, does not generalize to all forms of quiescence, such as Mek inhibition and CDK4/6 inhibition, likely because these are pharmacologically induced. It also does not generalize to all cell types, such as liposarcoma cells (Fig. 6D, vi), likely because of the vast differences in metabolism between fibroblasts and adipocytes (47).

One of the advantages of this flux-based approach is its ability to accommodate multiple types of omics data. Using transcriptomics, translomics, and proteomics data as inputs, external validation accurately predicted quiescent cells to have a deeper quiescence than proliferating cells (Fig. 6D i, iii, iv). Features in the model largely aligned with metabolic trends highlighted in Fig. 2: reactions corresponding to heme metabolism, carnitine shuttle, and fatty acid oxidation were enriched in the positive coefficients (up with quiescence deepening) while reactions associated with fatty acid biosynthesis were overrepresented in the negative coefficients (down with quiescence deepening) (Fig. 6B). The full set of features is listed in Table S2.

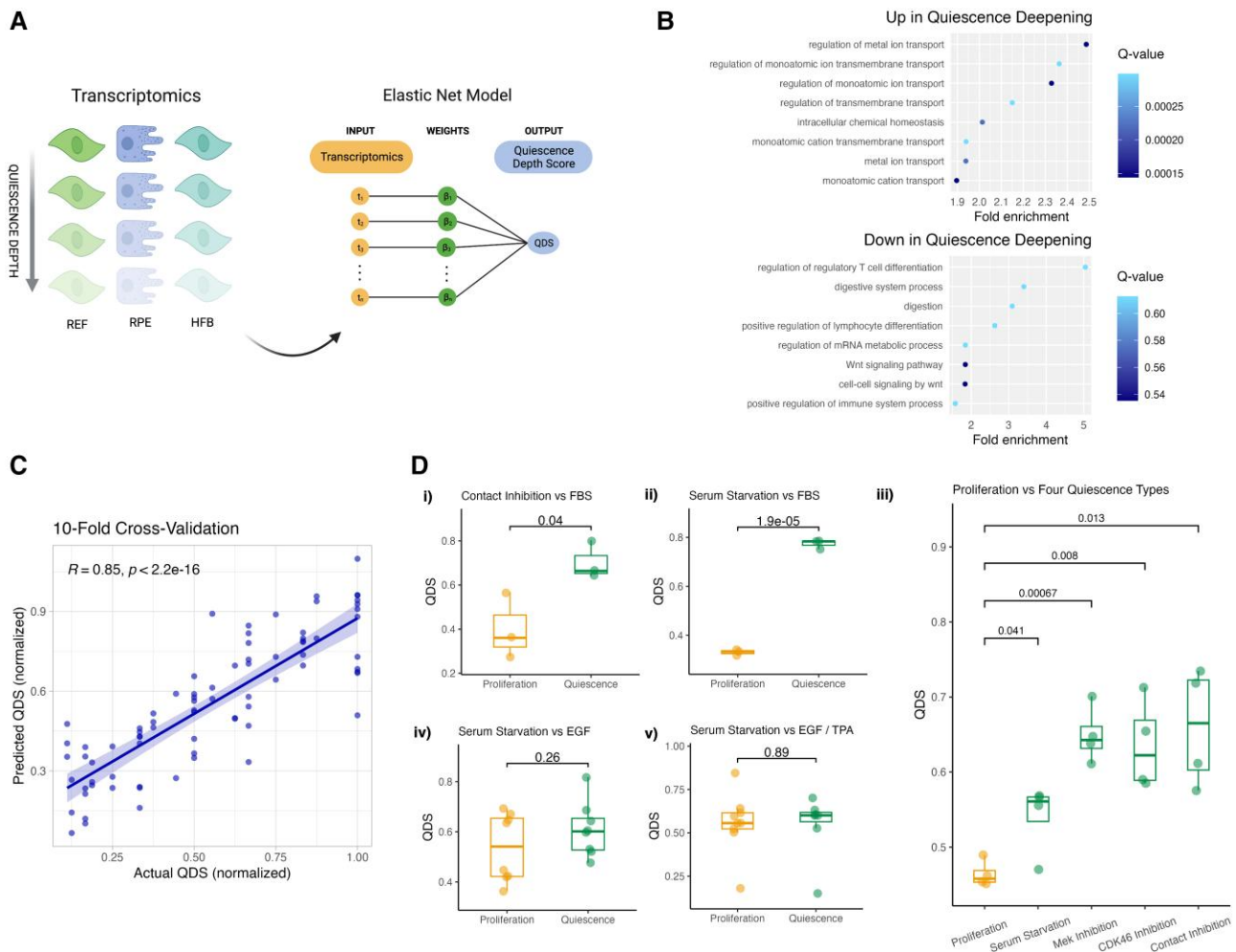


Fig. 7. Multi-cell type machine learning model accurately predicts QDS in a wide range of contexts. A) Diagram of the machine learning model. Transcriptomics from three datasets measuring quiescence deepening is inputted, and the model is trained to predict QDS (days since serum starvation, normalized to the maximum of each dataset) using elastic net regression. B) GO pathway enrichment analysis for positive and negative coefficients of model. The top eight biological processes for negative coefficients (down in quiescence deepening, Q -value < 0.05) and positive coefficients (up in quiescence deepening, Q -value < 0.05) as ranked by Q -value of enrichment. C) K -fold cross-validation accuracy plot. The model is trained and optimized on 90% of the data and tested in the remaining 10% for 10 random subsets of the data. D) External validation: model predictions for proliferating vs. quiescent cells in various cell types and experimental conditions. (i) Proliferation is induced using FBS. Quiescence induced with 7-day contact-inhibition. Comparisons were performed with a two-sample t test. Source: GSE117444 (45). (ii) Proliferation is induced using FBS. Quiescence induced with serum starvation. Comparisons were performed with a two-sample t test. Source: GSE74620 (46). (iii) Proliferation is induced using EGF, horse serum, insulin, and hydrocortisone. Quiescence is induced with serum starvation, Mek inhibition, CDK46 inhibition, and contact inhibition. Comparisons were performed with a two-sample t test. Source: GSE122927 (41). (iv) QDS for proliferating vs. quiescent cells. Proliferation is induced using EGF. Quiescence induced with serum starvation. Comparisons were performed with a two-sample t test. Source: GSE112295 (44). (v) Proliferation is induced using EGF or TPA. Quiescence induced with serum starvation. Comparisons were performed with a two-sample t test. Source: Ref. (42).

Multi-cell type quiescence depth predictor accurately predicts quiescence depth in a wide range of contexts

While the flux-based model provides both mechanistic interpretability and flexibility in terms of data types that can be inputted, the large number of computations it requires limits its application in external datasets. Hence a second, complementary model that directly uses transcriptomics data to predict QDS was also developed to facilitate high-throughput screening of perturbations to modify quiescence depth and explore quiescence in large external datasets. To build the model, we compiled quiescence-deepening gene expression datasets from three experimental conditions including rat embryo fibroblasts (REF) (5), human retinal pigment epithelial cells (RPEs) (48), and human dermal fibroblasts (HFB) (48) (Fig. 7A). Taken together, these datasets

encompass multiple species (rat and human), multiple cell types (epithelial and fibroblast), and multiple quiescence induction methods (serum starvation and contact inhibition). Similar to before, quiescence depth was denoted as the days since quiescence induction began, and each quiescence depth was normalized to the maximum of each dataset. Transcriptomics gene expression was set as the input to the model and the output was set as the normalized quiescence depth. Like before, an elastic net architecture was chosen to balance underfitting and overfitting the model.

As before, the model's accuracy was validated internally and externally. For k -fold cross-validation, the model accurately predicted quiescence depth across the entire range of values when trained on 90% of the data and tested on the final 10% (Fig. 7C). To validate the accuracy of the model in independent datasets,

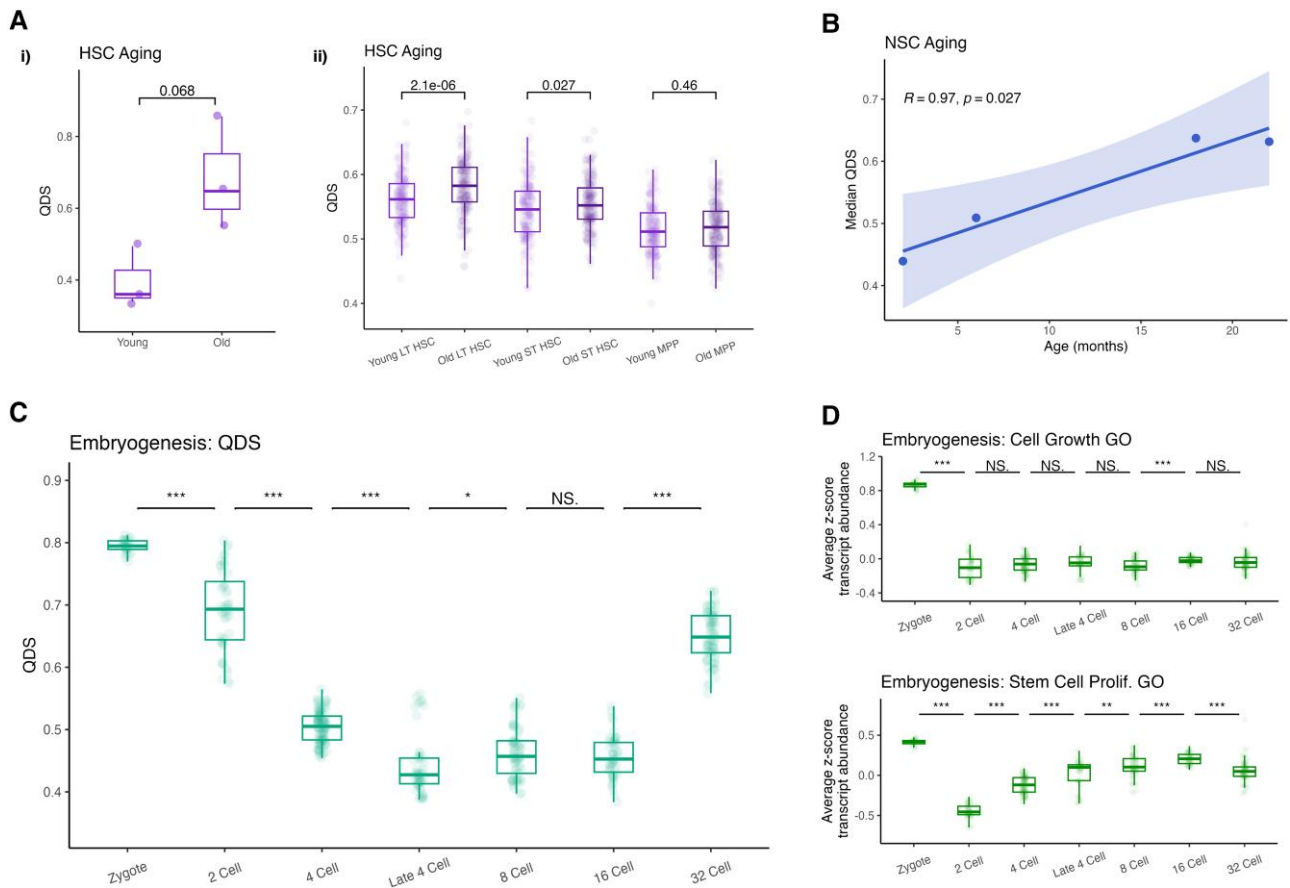


Fig. 8. Changes in quiescence depth are associated with stem cell aging and embryogenesis. A) HSC aging: (i) Older HSCs have a trend toward higher QDS than young HSCs. Source: GSE109546. (ii) Older LT HSCs and ST HSCs have higher QDS than young cells of the same type. No aging trend was observed in MPPs. Single-cell RNA-seq of aging HSCs. Source: GSE59114. B) Median QDS and age correlate in mouse subventricular zone (SVZ) neural stem cell (NSC) niche (Pearson correlation). Source: GSE104651. C) During embryogenesis, quiescence depth is predicted to fall and then rise between the zygote and 32-cell state. Single cell RNA-seq measured for preimplantation mouse blastomeres. Source: GSE136714. D) GO transcriptomic analysis of proliferation markers. Shown here are the mean z-score-normalized transcript levels for transcripts belonging to GO cell growth and GO stem cell proliferation. While GO cell growth-related genes show no definitive pattern, stem cell proliferation-related genes rise and then fall in a pattern that (except for the zygote stage) aligns with the fall and rise in QDS seen in early embryogenesis.

we examined five external gene expression datasets involving eight distinct experimental conditions for proliferating vs. quiescent cells (42–45). Strikingly, quiescent cells were accurately predicted to have a deeper quiescence than proliferating controls for six of the eight conditions tested. This was also observed across multiple types of growth stimulation and multiple types of quiescence induction, including contact inhibition, serum starvation, Mek inhibition, and CDK4/6 inhibition (Fig. 7D).

To gain a better understanding of the pathways captured in this model, we performed a gene ontology (GO) enrichment analysis for the negative features (down in quiescence deepening) and the positive features (up in quiescence deepening) (Fig. 7B). Various ion transport pathways are enriched in the positive model coefficients whereas several immune-related and Wnt-related pathways are enriched in the negative features. The full set of features is listed in Table S3.

Changes in quiescence depth are predicted during stem cell aging and embryogenesis

Fujimaki et al. (5) observed a deepening of quiescence during stem cell aging. Exploring this area with our multi-cell type machine learning model, we examined quiescence depth across age for both hematopoietic stem cells (HSCs) (49, 50) and NSCs (51). For

datasets with significant missing features, we reconstructed the model using genes shared between that dataset and the training gene set (see Methods).

For both NSCs and HSCs, we observed a trend reflecting a deepening of quiescence with age (Fig. 8A and B). For HSCs, we applied the model to a single-cell transcriptomics dataset measuring gene expression in long-term (LT) HSCs, short-term (ST) HSCs, and multipotent progenitor cells (MPPs). Intriguingly, we see a deepening of quiescence in LT and ST HSCs with age but not MPPs, suggesting aging is associated with cell-type-specific changes in quiescence depth. For NSCs, the trend of increasing quiescence depth followed by a reversal matches the nonmonotonic changes in NSC behavior with aging observed by Apostolopoulou et al. (51) (Fig. 8B).

Examining quiescence in a different biological context, we investigated quiescence depth during embryogenesis, specifically focusing on the zygote to 32-cell state. Here, the model predicted a fall and then a rise in quiescence depth between the zygote and 32-cell (Fig. 8C). This suggests that during embryogenesis, cells are most proliferative around the four-cell state and less so before and after. To corroborate these findings, we examined the normalized expression levels of genes belonging to two proliferation-related gene sets (Fig. 8D). While *cell growth* gene expression shows no definitive trends, *stem cell proliferation* gene expression (aside from the zygote stage) rises and then falls during

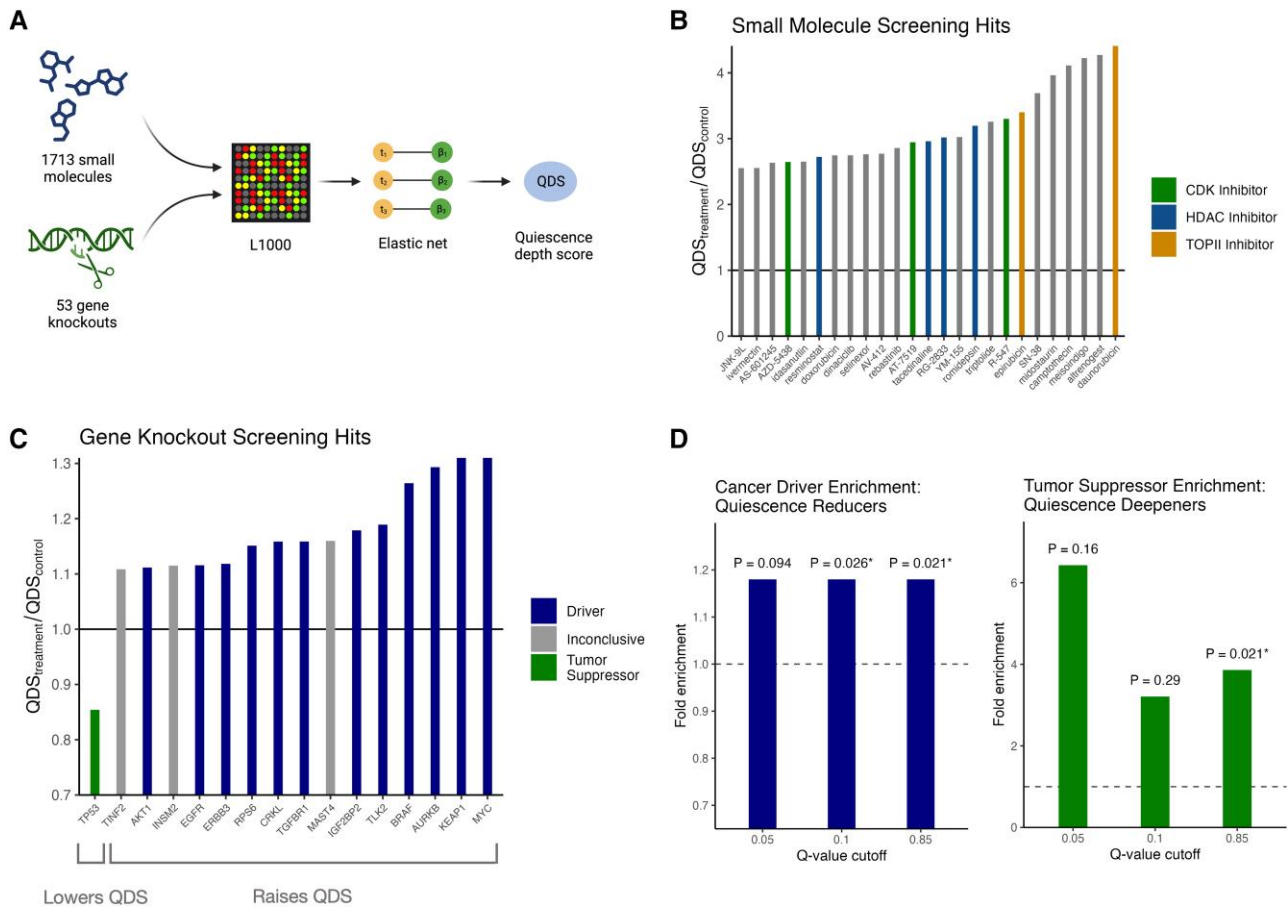


Fig. 9. Small molecules and gene knockouts predicted to alter quiescence align strongly with experimental results. A) Overview of perturbation screening: L1000 data measured as part of the library of integrated network-based cellular signatures program for two classes of perturbations (small molecules and gene knockouts) is inputted into the QDS predictor to screen perturbations for their effect on quiescence depth. B) Top small molecules whose knockout is predicted to alter quiescence depth. For each of the 1,713 small molecules tested, conditions are controlled to a 24-h timepoint and 10 μM dose across multiple cell lines. Top 25 predictions with Q-value < 0.05 shown here, as ranked by magnitude of quiescence depth score modulation. Q-value is the Benjamin–Hochberg-corrected P-value of a two-sample t test between predicted QDS for the control condition (drug vehicle) and treatment condition (small molecule compound). C) Top genes whose knockout is predicted to alter quiescence depth (Q-value < 0.05). D) Enrichment of tumor suppressors among quiescence deepeners (genes whose knockout lowers QDS and increases proliferative potential), and enrichment of oncogenes/cancer drivers among quiescence reducers (genes whose knockout raises QDS and decreases proliferative potential). P-value calculated using the hypergeometric function. Results are shown at three Q-value thresholds for selecting top gene knockout hits: 0.05, 0.1, and 0.85. Cancer annotations sourced from the CancerMine database. Genes with a majority of tumor suppressor annotations in CancerMine are classified as tumor suppressors, and genes with a majority of cancer driver or oncogene annotations are classified as cancer drivers.

early embryogenesis, which aligns with the changes in quiescence depth predicted by the machine learning model. Given that embryonic cells are a type of stem cell, the *stem cell proliferation* gene set serves as a more relevant validation of these predictions. Taken together, our model implies that quiescence depth regulation is a critical factor in both development and aging.

Various small molecules and genes predicted to alter quiescence align strongly with experimental results

We next applied the multi-cell type QDS predictor to screen a range of perturbations for their ability to modify quiescence depth. As input to the model, we used data measured as part of the library of integrated cellular signatures (LINCS) program that measures gene expression in response to small molecule and genetic perturbations using the L1000 assay (52) (Fig. 9A).

Importantly, we reconstructed the model using genes shared between the high confidence “landmark” L1000 gene set and the training gene set.

First, we screened all 1,713 small molecules in the LINCS dataset for their ability to alter quiescence depth. For each molecule, QDS was predicted and compared to QDS values predicted for a vehicle control with a two-sample t test. The ratio between the average QDS in the treatment group and the average QDS in the control group was then computed. At a false discovery rate (FDR) of 5%, 488 small molecules were predicted to raise or lower quiescence depth. The top 25 predictions are highlighted in Fig. 8B, and the full list of compounds is listed in Table S4.

Among the top compounds predicted to deepen quiescence, cyclin-dependent kinase (CDK), histone deacetylase (HDAC), and topoisomerase II (TOPII) inhibitors are common (Fig. 9B). In alignment with their predicted role as quiescence deepeners, many of the identified small molecules predicted to deepen quiescence

have direct experimental evidence for inducing quiescence through cell cycle arrest, including the CDK inhibitors R-547 (53), AT-7519 (54), and AZD-5438 (55). Others may be novel compounds to alter quiescence. Of the 488 small molecules predicted to alter quiescence, 51 are predicted to lower quiescence depth. While several of these compounds have antiproliferative activity and are thus likely false positives, the hormone estradiol is among the compounds predicted to lower quiescence, which aligns with the proliferation-promoting effects of estrogen (56).

Next, we screened all 53 gene knockouts in the LINCS dataset for their ability to modulate quiescence. QDS values predicted for each gene knockout were compared to QDS predicted for an empty gene vector control. As with the metabolic model knockouts, we used the CancerMine database to validate our predictions. Here, for both directions of QDS modulation, a striking percentage of genes were validated with experimental cancer data. At an FDR of 5%, of the genes with annotations, all genes whose knockout lowered QDS are annotated as tumor suppressors, and all genes whose knockout raised QDS are considered cancer drivers (Fig. 9C).

For the genes whose knockout is predicted to lower quiescence depth, the strongest prediction was made for tumor protein p53 (TP53), a gene famous for its tumor-suppressive and antiproliferative properties (57). At more lenient differential QDS thresholds, three of the top five genes whose knockout lowers QDS are tumor suppressors, an enrichment of 3.86-fold over the background gene set (P -value 0.021).

For genes whose knockout is predicted to deepen quiescence, results are similarly striking. At Q -value thresholds up to 0.3, all genes whose knockout is predicted to raise quiescence depth are considered oncogenes or cancer drivers. Among the top predictions are several notorious oncogenes including proto-oncogene C-Myc (MYC), which is dysregulated in over 50% of human cancers (58), and proto-oncogene B-Raf (BRAF), whose mutation is associated with nearly 7% of all human cancers (59). The top gene knockout hits for both directions of QDS alteration are shown in Fig. 9C along with their CancerMine annotation. The full list of genes is listed in Table S5.

Discussion

Here, we employed two distinct yet complementary approaches to identify modulators of quiescence depth: *in silico* gene knockouts in metabolic models, and a multi-cell type quiescence depth predictor that facilitates perturbation screening. In doing so, we also simulated metabolic and epigenetic changes underlying quiescence deepening, revealing substantial shifts in lipid metabolism and epigenetic modifications as quiescence shifts from a shallow to a deep state. For both screening approaches, top quiescence-altering candidates align strongly with experimental data and suggest various pathways capable of shifting quiescence toward a shallower or deeper state.

The most striking metabolic trends identified during quiescence deepening involve lipid metabolism, namely the fall in various kinds of lipid anabolism and the rise of lipid catabolism, particularly fatty acid oxidation. In alignment with these results, fatty acid oxidation is a key hallmark of quiescence across multiple cell types. Comparing the metabolomics of proliferating vs. quiescent fibroblasts, Yao et al. (20) observed that quiescent fibroblasts showed higher levels of carnitine-shuttle-dependent fatty acid oxidation. Further, they observed lower rates of complex-lipid biosynthesis in quiescent cells. Additional studies have shown that in both fibroblasts (60) and NSCs (61), quiescent cells are so reliant upon fatty acid oxidation for energy production

that inhibiting the pathway impairs cell viability. Both lend experimental evidence to a metabolic shift from fatty acid anabolism to catabolism as cells transition from proliferation to progressively deeper stages of quiescence.

Based on epigenetic simulations, similarly strong trends were identified in histone modifications with acetylation falling and methylation rising as quiescence deepens. Quiescence is typically associated with reduced gene expression and more compact chromatin architecture. It is also associated with changes in histone methylation, although the direction of this change is variable and depends on the specific histone residue (26). For acetylation, a clearer pattern is evident. In agreement with the simulated results, quiescence induction is associated with a reduction in histone acetylation across multiple cell types and histone residues. Since acetylation promotes a more accessible chromatin structure, this reduction—in part—mediates the transition to lower levels of gene expression during the transition from proliferation to quiescence (26). An investigation into the metabolic pathways driving the trends in histone modifications revealed that shifts in fatty acid metabolism—higher fatty acid oxidation, lower fatty acid biosynthesis—in quiescence could explain the observed trends and highlight the strong interplay between metabolism and epigenetics. Through our modeling, we showed how quiescence-inducing perturbations such as serum starvation can result in metabolic changes that influence epigenetic modifications, which alter gene expression and ultimately result in withdrawal from the cell cycle.

Given the alteration in quiescence depth observed in aging and disease (5, 60, 62), identifying chemical and genetic modulators of quiescence may be key to addressing these conditions. In this study, we identified potential modulators of quiescence depth using two approaches. Using rMTA, we identified genes whose knockout could robustly shift the metabolic state of a cell to a deeper or shallower state of quiescence. Particularly for quiescence reducers, results align well with experimental data. Notably, several genes predicted to affect quiescence depth here are involved in lipid metabolism, aligning with the shifts in lipid metabolism predicted in quiescence deepening.

We also developed two complementary quiescence depth predictors: a flux-based predictor that captures metabolic variations in quiescence deepening and a gene expression-based predictor based on multiple cell types and quiescence induction methods. Both models accurately predicted quiescent cells to have higher quiescence depth than proliferating cells for multiple types of growth induction, multiple types of quiescence induction, and multiple types of cells, suggesting they are generalizable to contexts beyond the datasets they were trained. Applying the multi-cell type model to screen L1000 perturbations generated predictions that align well with experimental data for both small molecules and gene knockouts. It also predicted changes in proliferative ability during embryogenesis, suggesting quiescence depth is an important biological factor from the earliest stages of development onto later aging processes. Interestingly, several small molecules whose knockout is predicted to affect quiescence depth target histone deacetylases (HDACs), underscoring the link between epigenetics and quiescence deepening, and connecting the metabolic modeling and knockout screening.

Several limitations to this study should be noted. Metabolic simulations are a powerful hypothesis-generating tool, and although several predictions align with experimental datasets, more comprehensive experimental validation is necessary to confirm their accuracy. Moreover, while GEMs are a well-established tool to study metabolism, their predictions are limited to the genes and metabolic pathways that are well described

in literature, which represents only a subset of mammalian metabolism. The small molecule and genetic perturbations will also require further validation to confirm their ability to alter quiescence depth.

In sum, we leverage two approaches to identify perturbations with strong potential to alter quiescence depth and with relevance to cancer and aging. We also show that GEMs can capture known aspects of quiescence deepening across metabolic and epigenetic levels and simulate the interplay between metabolic networks and epigenetic modifications. Our multi-cell type quiescence depth predictor is generalizable to a wide range of contexts. This study underscores the power of metabolic modeling and machine learning to characterize and predict modulators of cell states.

Methods

Differential expression analysis

To identify differentially expressed genes found in omics datasets for integration into the metabolic model, we employed a fold change-based approach. For this approach, the ratio between two groups was computed and transcripts with a fold change above a certain threshold were considered upregulated and those with a fold change below a certain threshold downregulated. A fold change threshold of 1.00 was used unless otherwise noted.

Flux prediction

To simulate metabolic reaction fluxes, we used two complementary approaches: flux sampling and a linear version of the iMAT algorithm (18, 24). For both approaches, we used the Rat-GEM metabolic model (16), which was chosen to match the biology of the rat embryonic fibroblast cells in which the quiescence deepening transcriptomics was measured. This model contains 2,953 genes, 8,377 metabolites, and 13,028 reactions.

For flux sampling, quiescence depth-specific metabolic models were created for nine stages of quiescence depth with three samples each by incorporating highly and lowly expressed genes to constrain the models using INIT (17, 63). Highly expressed genes were defined as genes whose quantile-normalized, log₂-normalized value was at least the mean plus α times the standard deviation in at least β percent of the three quiescence depth replicates (Eq. 1). Lowly expressed genes were defined similarly (Eq. 2).

$$\text{Gene}_{\text{up}} \geq \sum(\text{Log}_2(\text{Gene}_{\text{abun.}} + 1) \geq (\text{mean} + \alpha * \text{st. dev.})) \geq \beta * 3. \quad (1)$$

$$\text{Gene}_{\text{down}} \geq \sum(\text{Log}_2(\text{Gene}_{\text{abun.}} + 1) \leq (\text{mean} - \alpha * \text{st. dev.})) \geq \beta * 3. \quad (2)$$

α was chosen as the minimum value that generated metabolic models that could obtain a feasible flux solution (described in the next step) through flux sampling for all nine quiescence depths. α was raised in increments of 0.1 until a value of 1.1 was determined. β was chosen as 0.6 (meaning genes had to meet the specified criteria in at least two of the three replicates) to balance capturing changes in transcripts over time and still imposing a robust threshold for selecting up and down genes.

We then used coordinate hit-and-run with rounding (CHRR) (64) to perform flux sampling on the quiescence-specific metabolic models for each stage of quiescence depth. 1,000 flux samples were generated for each quiescence depth. The Gurobi

mathematical solver (version 9.5, Linux) was used to solve for the reaction fluxes on a high-performance computing cluster (65).

While flux sampling has advantages, including the absence of a presupposed objective function and a more comprehensive accounting of alternate flux solutions, it also takes considerable computational resources and time. Because of this, we also simulated reaction fluxes with a simpler approach using a linear version of the iMAT algorithm and compared the results to flux sampling (Fig. S4). This approach assumes a steady-state metabolism and solves for the reaction fluxes by optimizing an objective function (66). Here, reactions associated with upregulated genes are assigned higher flux values while reactions associated with downregulated genes are assigned lower values. Up and down genes are determined using the fold change approach described above based on differential expression compared to the lowest quiescence depth (QDS = 2); hence, eight quiescence depth scores were used with three replicates each (since the first quiescence depth was used as the reference). The Gurobi mathematical solver (version 9.0, macOS) was used here to solve for the reaction fluxes (65). Based on a subsystem analysis, we found that the predicted metabolic changes during quiescence deepening were largely similar between the linear iMAT algorithm and flux sampling (Fig. S4).

The linear version of the iMAT algorithm included three input parameters: κ (weight for downregulated genes), ρ (weight for upregulated genes), and ϵ (minimum absolute value for upregulated flux). To determine which parameter values to use, different ranges of values were tested to determine which parameter set minimized three-fold cross-validated mean squared error in predicting quiescence depth through elastic net regression. The same approach was used to choose the objective function to optimize for during flux balance analysis (FBA) (biomass, adenosine triphosphate [ATP], nicotinamide adenine dinucleotide [NADH], and glutathione were tested, and biomass was chosen) and the optimal fold change threshold for selecting upregulated and downregulated genes from the quiescence deepening transcriptomics to input into iMAT (1.008695652). The bounds of the fold change threshold optimization range were set close to 1 (1 to 1.2) to maximize the impact of genes on the fluxes and separate conditions in external datasets.

Histone methylation and acetylation simulation

To simulate histone methylation and acetylation, reactions were added to the rat-GEM metabolic model representing these histone modifications as described in Fig. 4A (24, 25). This simulation modeled the substrate metabolite abundance to predict histone modification flux. Then the objective value for this reaction was modified to a nonzero value and flux predictions were made for each QDS using the linear version of the iMAT algorithm described above. Linear iMAT was chosen over flux sampling as quiescence-associated fluxes were similar between flux sampling and iMAT and because this approach requires an objective function value for the histone-modification reaction to be a nonzero value (since the histone-modification reaction does not contribute to biomass optimization). The exact objective function value for the histone-modification reaction was determined by identifying the range of values that separated the histone-modification fluxes across QDS (i.e. fluxes that were not all identical) and resulted in biomass flux values that were on the same order of magnitude or higher than the flux through the histone-modification reaction (meaning the histone objective was not overoptimized relative to biomass). Figure S5 shows histone-modification fluxes and

biomass fluxes meeting these criteria for multiple objective function values.

Subsystem enrichment analysis

To perform the metabolic subsystem enrichment analysis, the proportion of reactions belonging to a particular subsystem within the highly correlating-with-quiescence depth reactions was divided by the proportion of reactions associated with that particular subsystem out of the total set of reactions. This was done separately for the positively correlating and negatively correlating reaction sets. The hypergeometric function was used to calculate the *P*-value for the enrichment of each subsystem. These *P*-values were corrected with the Benjamini–Hochberg FDR adjustment.

rMTA gene knockout screening

For the rMTA gene knockout screening (29, 64), we generated metabolic models for shallow quiescence (QDS of 2 to 3 days) and deep quiescence (QDS of 12 to 14 days) using INIT and selecting up and down genes as described above in the Flux prediction section. We performed two rounds of rMTA: one in which the source state was shallow quiescence and the target state was deep quiescence and another in which the source state was deep quiescence and the target state was shallow quiescence. For each round, the following procedure was used:

1. For the source state, we generated fluxes through flux sampling with CHRR using the same parameters described in the Flux prediction section (30, 64).
2. Using differential transcriptomics, we determined the transcripts that changed between the source state and target state at an FDR value of 0.05 (Benjamini–Hochberg-corrected *P*-value of two-sample *t* test between the source and target state).
3. We mapped these changes in transcripts to changes in metabolic reactions using gene-reaction mapping in the metabolic model.
4. We knocked out each gene in the metabolic model and determined its potential to shift these metabolic reactions from the source to the target state.

The rMTA screening was performed using the Gurobi mathematical solver (version 9.5, Linux) with default parameters on a high-performance computing cluster (65). Gene knockout predictions were validated using CancerMine as described below in *CancerMine validation*.

CancerMine validation

For the CancerMine database approach, gene annotation as a tumor suppressor, oncogene, or cancer driver was used to validate predictions (30).

To validate genes whose knockdown was predicted to lower quiescence depth (i.e. raise proliferative potential), the following criteria were used:

1. In the CancerMine database, a majority of annotations for this gene were tumor suppressors (genes whose downregulation increases cancer growth)

To validate genes whose knockdown was predicted to raise quiescence depth (i.e. lower proliferative potential), the following criteria were used:

1. In the CancerMine database, a majority of annotations for this gene were either oncogenes or cancer drivers (genes whose downregulation lowers cancer growth)

If no annotations were present or if annotations were split 50% tumor suppressor, 50% oncogene, or driver, validation for that gene was considered inconclusive.

Machine learning model creation and optimization

To build the flux-based QDS model, all GEM reaction fluxes generated by the linear version of iMAT were set as the input, and the output was set as QDS (days since serum starvation) without normalization to [0, 1]. The linear version of iMAT was chosen as the input because its predictions aligned well with flux sampling (see Fig. S4) while using considerably fewer computational resources. An elastic net architecture was chosen to strike a balance between underfitting and overfitting the reaction coefficients. Underfitting would impair the accuracy of the model in the training dataset and overfitting would hinder the generalizability of the model in external datasets, both outcomes we sought to avoid. The elastic net architecture accomplishes this by minimizing model error while also minimizing both the sum of the magnitude of coefficients and the sum of the square of coefficients (so-called L1 and L2 penalties, also known as lasso and ridge penalties). This ensures that only the most salient and generalizable features are incorporated into the model.

The model was implemented in MATLAB with the lasso function and accepted two key parameters: lambda, the regularization coefficients, and alpha, the ratio between L1 and L2 optimization (or lasso and ridge optimization). Both alpha and lambda were optimized to minimize three-fold cross-validated mean squared error in predicting QDS. The fold change threshold, objective function, and inputs to the linear version of iMAT were optimized similarly, as described in the Flux prediction section (Table S6). In total, 838 metabolic reactions were selected by elastic net regression in this model (i.e. they had nonzero coefficient values) out of 13,028 total reactions used in the training set.

To ensure the flux-based model was robust to the data integration method, we predicted quiescence depth across the entire range of values for fluxes generated via INIT and flux sampling (Fig. S7) as described in the Flux prediction section. Here, the model accurately predicted a deepening of quiescence for these fluxes despite vastly different transcriptomics incorporation methods.

To build the multi-cell type transcriptomics QDS predictor, we aggregated three transcriptomics datasets capturing quiescence deepening over time (5, 48, 67). Each dataset was converted to transcripts per million (TPM) and quantile-normalized to align the distributions between datasets (Fig. S6). A count filter was applied to omit any transcripts with a coverage of <10. Proliferating controls were also omitted. Then, principal component analysis was performed to identify outlier conditions, and any outliers were removed (Fig. S6). Finally, each transcript was z-score normalized across conditions. This resulted in 79 total samples across the three datasets. Normalized gene expression was set as the input to the model and QDS was set as the output.

For each sample, QDS was defined as the days since quiescence induction began. Quiescence depth was normalized to the maximum of each dataset, so all values were in the range of 0 to 1 (Fig. S6).

To construct the machine learning model, an elastic net regression model was built and optimized to predict QDS based on

transcript levels. As before, an elastic net architecture was chosen to strike a balance between underfitting and overfitting the model coefficients. The model was also implemented in MATLAB with the lasso function and accepted two key inputs: lambda, the regularization coefficients, and alpha, the ratio between ridge and lasso optimization. Both alpha and lambda were optimized to minimize 10-fold cross-validated mean squared error in predicting QDS (Table S7). In total, 1,306 genes were selected by elastic net regression in this model out of 9,304 genes in the training set.

To evaluate both models internally, we performed 10-fold cross-validation in which a random 90% of the data was used to train and validate the model, and the remaining 10% of the data was used as a test set to evaluate the model's accuracy. The trained and optimized model had never seen this last 10%. This process was repeated 10 times until all samples had been in the test group.

Enrichment analysis of transcriptomic elastic net model coefficients

To perform GO enrichment of the multi-cell type elastic net model QDS predictor, we used the R package *clusterProfiler* (68). We searched for GO biological processes that were enriched in the positive coefficients (up with quiescence deepening) and negative coefficients (down with quiescence deepening). For the background gene set, we used the list of genes used for training the elastic net model.

Application of models to external datasets

For external datasets, raw counts were filtered to omit genes with a count of <10, and each condition was size-normalized. For single-cell RNA-sequencing data, low coverage values were imputed using the algorithm Markov affinity-based graph imputation of cells (MAGIC) (69).

If fewer than 40% of model coefficient genes were found in a dataset after filtering, the elastic net model was reconstructed using genes shared between the training set and the external dataset using the same approach described in the Machine learning model creation and optimization section. This procedure was performed only for the dataset GSE109546.

L1000 perturbation screening

Level 5 (consensus z-score signature) L1000 gene expression data from the LINCS phase II (52) dataset was read into MATLAB using CMAP-associated scripts (70). Genes were filtered to include only the directly measured landmark gene set. The multi-cell type QDS predictor was reconstructed using only the genes common to the training set and the landmark gene set with the same approach described in the Machine learning model creation and optimization section.

Small molecule screening

One thousand seven hundred and thirteen small molecules were tested, and conditions were controlled to a 24-h post-small-molecule-administration timepoint and 10 μ M dose. Multiple cell lines were used for each small molecule. QDS predictions were made with the multi-cell type QDS predictor for each small molecule and compared to QDS predictions for a vehicle control with a two-sample t test. Small molecules with a significant difference in mean QDS between the treatment and control group were highlighted.

Gene knockout screening

Fifty-three gene knockouts were tested, and all conditions were measured 96 h postgene-knockout. Multiple cell lines were used for each gene knockout. QDS predictions were made with the multi-cell type QDS predictor for each gene and compared to QDS predictions for an empty gene vector control with a two-sample t test. Gene knockouts with a significant difference in the mean QDS between the treatment and control group were highlighted and validated as described in the CancerMine validation section.

Quantification and statistical analysis

Statistical tests were performed in MATLAB (version R2022A and R2022B) and R (version 4.3.0). All correlations were computed with the Pearson method. A two-tailed t test was used to determine the statistical significance of differences between groups based on the independence of observations in each sample, the normality of samples, similarity in sample variance, and random sampling of data. The sample size was determined based on the number of replicates in publicly available datasets and listed in figure captions for each comparison (no power analyses were performed). A paired t test was used to compare quiescence depth scores generated through the fold change approach (as these were paired samples). Enrichment significance was calculated using the hypergeometric test. *P*-values were corrected for multiple hypotheses through the Benjamini–Hochberg FDR correction. A *P*-value of 0.05 was used as a threshold for significance.

Supplementary Material

[Supplementary material](#) is available at PNAS Nexus online.

Funding

This work was supported by faculty start-up funds from the University of Michigan (UM), Camille and Henry Dreyfus Foundation, and the National Institute of General Medical Sciences (NIGMS) R35 GM13779501 to S.C.

Author Contributions

A.E. and S.C. conceived the study and wrote the manuscript. A.E. performed research with supervision by S.C.

Data Availability

All datasets used in this study were downloaded from public repositories. Transcriptomics data used to generate fluxes and train the quiescence depth predictors are available from GEO under accession numbers GSE124109 (5), GSE129964 (67), and GSE179848 (48). Omics data used to validate the models, explore quiescence in other contexts, and perform perturbation screening are available under the following GEO accession numbers: GSE117444 (45), GSE74620 (46), GSE122927 (41), GSE112295 (44), GSE109546 (49), GSE104651 (51), GSE136714 (71), GSE59114 (50), and GSE70138 (52). All code used to generate the results of this study is available at the following Github repository: <https://github.com/sriram-lab/quiescence-depth-modulation>.

References

- Coller HA. 2011. The essence of quiescence. *Science*. 334(6059): 1074–1075.

- 2 Coller HA, Sang L, Roberts JM. 2006. A new description of cellular quiescence. *PLoS Biol.* 4(3):e83.
- 3 Shea KL, et al. 2010. Sprouty1 regulates reversible quiescence of a self-renewing adult muscle stem cell pool during regeneration. *Cell Stem Cell.* 6(2):117–129.
- 4 Rumman M, Dhawan J, Kassem M. 2015. Concise review: quiescence in adult stem cells: biological significance and relevance to tissue regeneration. *Stem Cells.* 33(10):2903–2912.
- 5 Fujimaki K, et al. 2019. Graded regulation of cellular quiescence depth between proliferation and senescence by a lysosomal dimer switch. *Proc Natl Acad Sci U S A.* 116(45):22624–22634.
- 6 Gos M, et al. 2005. Cellular quiescence induced by contact inhibition or serum withdrawal in C3H10T1/2 cells. *Cell Prolif.* 38(2):107–116.
- 7 Bucher NLR. 1963. Regeneration of mammalian liver. *Int Rev Cytol.* 15:245–300.
- 8 Rodgers JT, et al. 2014. MTORC1 controls the adaptive transition of quiescent stem cells from G0 to GAlert. *Nature.* 510(7505):393–396.
- 9 Chell JM, Brand AH. 2010. Nutrition-responsive glia control exit of neural stem cells from quiescence. *Cell.* 143(7):1161–1173.
- 10 Lindell E, Zhong L, Zhang X. 2023. Quiescent cancer cells—a potential therapeutic target to overcome tumor resistance and relapse. *Int J Mol Sci.* 24(4):3762.
- 11 Mardinoglu A, Gatto F, Nielsen J. 2013. Genome-scale modeling of human metabolism—a systems biology approach. *Biotechnol J.* 8(9):985–996.
- 12 Price ND, Reed JL, Palsson BØ. 2004. Genome-scale models of microbial cells: evaluating the consequences of constraints. *Nat Rev Microbiol.* 2(11):886–897.
- 13 Hyduke DR, Lewis NE, Palsson BØ. 2013. Analysis of omics data with genome-scale models of metabolism. *Mol Biosyst.* 9(2):167–174.
- 14 Frezza C, et al. 2011. Haem oxygenase is synthetically lethal with the tumour suppressor fumarate hydratase. *Nature.* 477(7363):225–228.
- 15 Kutay M, Gozuacik D, Çaklr T. 2022. Cancer recurrence and omics: metabolic signatures of cancer dormancy revealed by transcriptome mapping of genome-scale networks. *OMICS.* 26(5):270–279.
- 16 Wang H, et al. 2021. Genome-scale metabolic network reconstruction of model animals as a platform for translational research. *Proc Natl Acad Sci U S A.* 118(30):e2102344118.
- 17 Agren R, et al. 2012. Reconstruction of genome-scale active metabolic networks for 69 human cell types and 16 cancer types using INIT. *PLoS Comput Biol.* 8(5):e1002518.
- 18 Zur H, Ruppig E, Shlomi T. 2010. iMAT: an integrative metabolic analysis tool. *Bioinformatics.* 26(24):3140–3142.
- 19 Rowe E, Palsson BO, King ZA. 2018. Escher-FBA: a web application for interactive flux balance analysis. *BMC Syst Biol.* 12(1):84.
- 20 Yao CH, et al. 2016. Exogenous fatty acids are the preferred source of membrane lipids in proliferating fibroblasts. *Cell Chem Biol.* 23(4):483–493.
- 21 Mentch SJ, et al. 2015. Histone methylation dynamics and gene regulation occur through the sensing of one-carbon metabolism. *Cell Metab.* 22(5):861–873.
- 22 Wellen KE, et al. 2009. ATP-citrate lyase links cellular metabolism to histone acetylation. *Science.* 324(5930):1076–1080.
- 23 Campit SE, Meliki A, Youngson NA, Chandrasekaran S. 2020. Nutrient sensing by histone marks: reading the metabolic histone code using tracing, omics, and modeling. *Bioessays.* 42(9):e2000083.
- 24 Shen F, Boccutto L, Pauly R, Srikanth S, Chandrasekaran S. 2019. Genome-scale network model of metabolism and histone acetylation reveals metabolic dependencies of histone deacetylase inhibitors. *Genome Biol.* 20(1):49.
- 25 Chandrasekaran S, et al. 2017. Comprehensive mapping of pluripotent stem cell metabolism using dynamic genome-scale network modeling. *Cell Rep.* 21(10):2965–2977.
- 26 Bonitto K, Sarathy K, Atai K, Mitra M, Coller HA. 2021. Is there a histone code for cellular quiescence? *Front Cell Dev Biol.* 9:739780.
- 27 McDonnell E, et al. 2016. Lipids reprogram metabolism to become a major carbon source for histone acetylation. *Cell Rep.* 17(6):1463–1472.
- 28 Izzo LT, et al. 2023. Acetylcarnitine shuttling links mitochondrial metabolism to histone acetylation and lipogenesis. *Sci Adv.* 9(18):eadf0115.
- 29 Valcárcel LV, Torrano V, Tobalina L, Carracedo A, Planes FJ. 2019. RMTA: robust metabolic transformation analysis. *Bioinformatics.* 35(21):4350–4355.
- 30 Lever J, Zhao EY, Grewal J, Jones MR, Jones SJM. 2019. CancerMine: a literature-mined resource for drivers, oncogenes and tumor suppressors in cancer. *Nat Methods.* 16(6):505–507.
- 31 Kwon JS, et al. 2017. Controlling depth of cellular quiescence by an Rb-E2F network switch. *Cell Rep.* 20(13):3223–3235.
- 32 Nevins JR. 2001. The Rb/E2F pathway and cancer. *Hum Mol Genet.* 10(7):699–703.
- 33 Schroeder B, et al. 2021. Fatty acid synthase (FASN) regulates the mitochondrial priming of cancer cells. *Cell Death Dis.* 12(11):977.
- 34 Zhang Y, et al. 2022. ELOVLs predict distinct prognosis value and immunotherapy efficacy in patients with hepatocellular carcinoma. *Front Oncol.* 12:884066.
- 35 Ershov P, et al. 2021. Enzymes in the cholesterol synthesis pathway: interactomics in the cancer context. *Biomedicines.* 9(8):895.
- 36 Petrelli F, et al. 2023. Mitochondrial pyruvate metabolism regulates the activation of quiescent adult neural stem cells. *Sci Adv.* 9(9):eadd5220.
- 37 Shea HC, Head DD, Setchell KDR, Russell DW. 2007. Analysis of HSD3B7 knockout mice reveals that a 3 α -hydroxyl stereochemistry is required for bile acid function. *Proc Natl Acad Sci U S A.* 104(28):11526–11533.
- 38 Gurbanov R, Tunçer S. 2018. PNP (purine nucleoside phosphorylase). *Atlas Genet Cytogenet Oncol Haematol.* 22:1–2.
- 39 Fitzgibbon J, et al. 1996. Human guanylate kinase (GUK1): cDNA sequence, expression and chromosomal localisation. *FEBS Lett.* 385(3):185–188.
- 40 Flinn AM, Gennery AR. 2018. Adenosine deaminase deficiency: a review. *Orphanet J Rare Dis.* 13(1):65.
- 41 Min M, Spencer SL. 2019. Spontaneously slow-cycling subpopulations of human cells originate from activation of stress-response pathways. *PLoS Biol.* 17(3):e3000178.
- 42 Sharma KL, et al. 2021. Mitogen-induced transcriptional programming in human fibroblasts. *Gene.* 800:145842.
- 43 Lee HJ, et al. 2017. Proteomic and metabolomic characterization of a mammalian cellular transition from quiescence to proliferation. *Cell Rep.* 20(3):721–736.
- 44 Clamer M, et al. 2018. Active ribosome profiling with RiboLace. *Cell Rep.* 25(4):1097–1108.e5.
- 45 Mitra M, et al. 2018. Alternative polyadenylation factors link cell cycle to migration. *Genome Biol.* 19(1):176.
- 46 Kovatcheva M, et al. 2017. ATRX is a regulator of therapy induced senescence in human cells. *Nat Commun.* 8(1):386.
- 47 Ali AT, Hochfeld WE, Myburgh R, Pepper MS. 2013. Adipocyte and adipogenesis. *Eur J Cell Biol.* 92(6–7):229–236.

- 48 Sturm G, et al. 2023. Oxphos defects cause hypermetabolism and reduce lifespan in cells and in patients with mitochondrial diseases. *Commun Biol.* 6(1):22.
- 49 Maryanovich M, et al. 2018. Adrenergic nerve degeneration in bone marrow drives aging of the hematopoietic stem cell niche. *Nat Med.* 24(6):782–791.
- 50 Kowalczyk MS, et al. 2015. Single-cell RNA-seq reveals changes in cell cycle and differentiation programs upon aging of hematopoietic stem cells. *Genome Res.* 25(12):1860–1872.
- 51 Apostolopoulou M, et al. 2017. Non-monotonic changes in progenitor cell behavior and gene expression during aging of the adult V-SVZ neural stem cell niche. *Stem Cell Rep.* 9(6):1931–1947.
- 52 Subramanian A, et al. 2017. A next generation connectivity map: L1000 platform and the first 1,000,000 profiles. *Cell.* 171(6):1437–1452.e17.
- 53 Hacıoğlu B, Kuş G, Kutlu HM, Kabadere S. 2020. The effect of R547, a cyclin-dependent kinase inhibitor, on hepatocellular carcinoma cell death. *Turk J Biol.* 44(1):24–33.
- 54 Zhao H, et al. 2019. Study of the mechanism by which dinaciclib induces apoptosis and cell cycle arrest of lymphoma Raji cells through a CDK1-involved pathway. *Cancer Med.* 8(9):4348–4358.
- 55 Raghavan P, et al. 2012. AZD5438, an inhibitor of Cdk1, 2, and 9, enhances the radiosensitivity of non-small cell lung carcinoma cells. *Int J Radiat Oncol Biol Phys.* 84(4):e507–e514.
- 56 Tian JM, Ran B, Zhang CL, Yan DM, Li XH. 2018. Estrogen and progesterone promote breast cancer cell proliferation by inducing cyclin G1 expression. *Braz J Med Biol Res.* 51(3):1–7.
- 57 Mantovani F, Collavin L, Del Sal G. 2019. Mutant p53 as a guardian of the cancer cell. *Cell Death Differ.* 26(2):199–212.
- 58 Chen H, Liu H, Qing G. 2018. Targeting oncogenic Myc as a strategy for cancer treatment. *Signal Transduct Target Ther.* 3:5.
- 59 Śmiech M, Leszczyński P, Kono H, Wardell C, Taniguchi H. 2020. Emerging braf mutations in cancer progression and their possible effects on transcriptional networks. *Genes (Basel).* 11(11):1342.
- 60 Ortmayr K, Zampieri M. 2022. Sorting-free metabolic profiling uncovers the vulnerability of fatty acid β -oxidation in in vitro quiescence models. *Mol Syst Biol.* 18(9):e10716.
- 61 Knobloch M, et al. 2017. A fatty acid oxidation-dependent metabolic shift regulates adult neural stem cell activity. *Cell Rep.* 20(9):2144–2155.
- 62 Brunet A, Goodell MA, Rando TA. 2023. Ageing and rejuvenation of tissue stem cells and their niches. *Nat Rev Mol Cell Biol.* 24(1):45–62.
- 63 Heirendt L, et al. 2019. Creation and analysis of biochemical constraint-based models using the COBRA Toolbox v3.0. *Nat Protoc.* 14(3):639–702.
- 64 Haraldsdóttir HS, Cousins B, Thiele I, Fleming RMT, Vempala S. 2017. CHRR: coordinate hit-and-run with rounding for uniform sampling of constraint-based models. *Bioinformatics.* 33(11):1741–1743.
- 65 Gurobi Optimization, LLC. 2023. Gurobi optimizer. <https://www.gurobi.com>.
- 66 Orth JD, Thiele I, Palsson BØ. 2010. What is flux balance analysis? *Nat Biotechnol.* 28(3):245–248.
- 67 Rajapakse D, et al. 2020. Amelotin is expressed in retinal pigment epithelium and localizes to hydroxyapatite deposits in dry age-related macular degeneration. *Transl Res.* 219:45–62.
- 68 Wu T, et al. 2021. clusterProfiler 4.0: a universal enrichment tool for interpreting omics data. *Innovation (Camb).* 2(3):100141.
- 69 van Dijk D, et al. 2018. Recovering gene interactions from single-cell data using data diffusion. *Cell.* 174(3):716–729.e27.
- 70 Enache OM, et al. 2019. The GCTx format and cmap{Py, R, M, J} packages: resources for optimized storage and integrated traversal of annotated dense matrices. *Bioinformatics.* 35(8):1427–1429.
- 71 Wang Y, et al. 2021. Single-cell multiomics sequencing reveals the functional regulatory landscape of early embryos. *Nat Commun.* 12(1):1247.