science.sciencemag.org/cgi/content/full/science.aax6234/DC1

Supplementary Materials for

Massively multiplex chemical transcriptomics at single-cell resolution

Sanjay R. Srivatsan*, José L. McFaline-Figueroa*, Vijay Ramani*, Lauren Saunders, Junyue Cao, Jonathan Packer, Hannah A. Pliner, Dana L. Jackson, Riza M. Daza, Lena Christiansen, Fan Zhang, Frank Steemers, Jay Shendure†‡, Cole Trapnell†‡

*These authors contributed equally to this work.
†These authors contributed equally to this work.
‡Corresponding author. Email: shendure@uw.edu (J.S.); coletrap@uw.edu (C.T.)

Published 5 December 2019 on *Science* First Release
DOI: 10.1126/science.aax6234

This PDF file includes:

Other Supplementary Material for this manuscript includes the following:
(available at science.sciencemag.org/cgi/content/full/science.aax6234/DC1)

Materials and Methods

Cell Culture

A549 cells and K562 cells were a kind gift from Dr. Robert Bradley (UW) and Dr. David Hawkins (UW), respectively. MCF7 (cat no. HTB-22), NIH3T3 (cat no. CRL-1658) and HEK293T (cat no. CRL-11268) cells were purchased from ATCC. A549 and MCF7 cells were cultured in DMEM (ThermoFisher, 11995073) media supplemented with 10% FBS (ThermoFisher, cat no. 26140079) and 1% penicillin-streptomycin (ThermoFisher, 15140122). K562 cells were cultured in RPMI 1640 (Fisher Scientific, cat no. 11-875-119) supplemented with 10% FBS and 1% penicillin-streptomycin and maintained between 0.2-1 x 10^6 cells/ml. All cells were cultured at 37C with 5% CO2. Adherents cells were split when they reached 90% confluence by washing with DPBS (Life Technologies, cat no. 14190-250), trypsinizing using TryPLE (Fisher Scientific, cat no. 12-604-039) and split at either 1:4 (MCF7) or 1:10 (A549, NIH3T3 and HEK293T).

Compound Preparation

Dexamethasone was purchased from Sigma-Aldrich and resuspended in molecular biology grade ethanol (Fisher Scientific). BMS-345541 (S8044), Vorinostat (S1047), and Nutlin-3a (S8059) were acquired from Selleck Chemicals and resuspended in DMSO (VWR Scientific, 97063-136). Cherry-picked 96-well compound screens were acquired from Selleck Chemicals resuspended to 10 mM in DMSO (**Table S3**). Compounds were diluted in their respective vehicle to 1000x of their desired treatment concentration and stored at -80C until use.

Drug treatment

For 96-well experiments, adherent cells were trypsinized, washed with PBS and plated in tissue culture treated 96 well flat bottom plates (Thermo Fisher Scientific, cat no. 12-656-66) at 25,000 cells per well in 100 µL of media. Suspension cells were washed with PBS and plated in 96 well V-bottom tissue culture plates (Thermo Fisher Scientific, cat no. 549935) at 25,000 cells per well in 100 µL of media. Cells were allowed to recover for 24 hours before treatment with 1 µL of a 1:10 dilution of the appropriate compound or vehicle in PBS to maintain a vehicle concentration of 0.1% for all wells. Cells were then exposed to small molecules at the specified concentration for either 24 or 72 hours. For experiments where cells were co-treated with HDAC inhibitors and either acetate, pyruvate, citrate, ACSS2 inhibitor (EMD Millipore Inc., Cat No. 533756,), ACLY inhibitor (Cayman Chemicals, BMS-303141 Cat No. 943962-47-8) or PDH inhibitor (Cayman Chemicals, Cat No. 504817), cells were treated 24 hours after plating and harvested after 24 hours. In this set of experiments, all wells contained a final concentration of 0.2% DMSO to match treatment with both the HDAC inhibitor and inhibitors of metabolic processes.

CellTiter Glo

A549, MCF7 and K562 cells were seeded in 96 well plates, allowed to attach for 24 hours and treated with BMS345541, dexamethasone, nutlin-3A, SAHA, as described above. 24 hours post treatment, plates were allowed to reach room temperature and viability estimated using the CellTiter-Glo viability assay (Promega) according to manufacturer's instructions. Luminescence was recorded using a BioTek synergy plate reader. For each drug treatment luminescence readings were normalized to the average luminescence intensities of vehicle DMSO treated wells.

Cell counts of bosutinib exposed cells

A549, MCF7 and K562 cells were seeded in 12 well plates at 2.8 x 10^5 cells per well. After 24 hours to allow for A549 and MCF7 attachment, cells were exposed for 24 hours to 0.1, 1 and 10 µM bosutinib or DMSO vehicle control. After treatment, adherent cells were detached using TrypLE or directly resuspended in 1 mL of media and cells counted on a Countess II FL automated cell counter (ThermoFisher).

Cancer cell line encyclopedia and connectivity map data and analysis

Pharmacological profiling data was downloaded from the Cancer cell line encyclopedia (CCLE) data portal (https://portals.broadinstitute.org/ccle/data). Data was the isolated and plotted for cell line of haematopoietic and lymphoid, lung and breast tissue origin exposed to the Abl inhibitors AZD0530 and nilotinib. Connectivity map (CMAP) data was downloaded from the CLUE command app in the CMAP data portal (https://clue.io/command?q=/home). Top connections and connectivity scores (obtained using the /conn command) were exported between the MEK inhibitor perturbagen class (CP_MEK_INHIBITOR) and HSP inhibitor perturbagen class (CP_HSP_INHIBITOR) across all cell lines (Summary) or individual cell lines that overlap with our study (A549 and MCF7). Results were then filtered for data from inhibitor exposure. To determine how connectivities change across all vs. individual cell lines, we filtered for the top connections that overlap with the connectivity summary in data from individual cell lines. Connectivity scores were subjected to a threshold value of 90 as in the associated CMAP study (*11*).

Flow cytometry

A549 and MCF7 cells were seeded in 6 cm dishes at 1.6 x 10^6 cells per plate. K562 cells were seeded in T25 cm2 flasks at 1.6 x 10^6 cells per flask. After 24 hours to allow for A549 and MCF7 attachment cells were exposed for 24 hours to 10 µM abexinostat, 10 µM pracinostat or DMSO as a vehicle control. After treatment cells were harvested as described above, pellets washed twice in PBS, resuspended in 500 µL of cold PBS and fixed by the addition of 5 mL of ice-cold ethanol while vortexing at low speed. Cells were stored at -20C prior to processing for flow cytometry analysis. For flow cytometry, ethanol was removed and fixed cells washed twice with PBS containing 1% BSA (PBS-B) and blocked for 1 hour at room temperature. Then, blocking buffer was removed and cells were incubated in PBS containing 1% BSA and 0.1% tryton X-100 (PBS-BT) as well as a 1:500 dilution of mouse anti-acetyl-lysine antibody (cat no. ICP0390, ImmuneChem Pharmaceuticals Inc) for 2 hours at room temperature. After incubation, cells were washed twice with PBS-BT and incubated with goat anti-mouse Alexa-647 in PBS-BT for 1 hour at room temperature. Lastly, cells were washed twice with PBS-BT, once with PBS-B and resuspended in PBS-B containing 5 µg/ml Hoechst 33258 (Life Sciences Technologies) to stain the DNA. Then the levels of total acetylated-lysine and DNA content was analyzed by flow cytometry on an LSRII flow cytometer (BD Biosciences). Quantification and downstream analysis was performed using FlowJo10 (FlowJo.LLC).

Cell harvest, nuclei isolation and sample hashing

For the harvest of adherent cells, media was removed, and cells were rinsed with 100 µL of DPBS and tryspinized with 50 µL of Tryp-LE for 15 minutes at 37C. Once cells had detached from the culture plate, the reaction was quenched with 150 µl of ice-cold DMEM containing 10% FBS. Cell suspensions were generated by pipetting and the entire volume was transferred to

a 96 well V-bottom plate. Cells were then pelleted by centrifugation at 300 x *g* for 6 minutes, washed with 100 µL of ice-cold DPBS and re-pelleted at 300 x *g* for 6 minutes.

Lysis was conducted in the 96 well V-bottom plate. Following removal of PBS, cell suspensions were lysed and labeled with 50 µL of cold lysis buffer (10 mM Tris-HCl, pH 7.4, 10 mM NaCl, 3 mM MgCl2, 0.1% IGEPAL CA-630) (*24*) supplemented with 1% Superase RNA Inhibitor and 400 femtomoles of hashing oligo of the form 5'-GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAG-[10bp-barcode]-BAAAAAAAAAAAAAAAAAAAAAAAAAAAAA-3' where B is G, C or T (IDT). For the large compound screen, 500 femtomoles of an additional oligo was used to uniquely index each 96 well treatment plate. After lysis with 3 strokes of multichannel pipette, cells were fixed by addition of 200 µL of fixation buffer (5% Paraformaldehyde, 1.25x PBS). Nuclei were then fixed on ice for 15 minutes before pooling into a trough. Nuclei were pooled by plate into a 50 mL conical tube and pelleted by centrifugation at 500 x *g* for 5 minutes. Subsequently, cells were resuspended in 500 µL of nuclei suspension buffer (NSB; (10 mM Tris-HCl, pH 7.4, 10 mM NaCl, 3 mM MgCl2, 1% Superase RNA Inhibitor, 1% 0.2mg/mL Ultrapure BSA)). Finally, nuclei from all plates were pooled into a single conical tube and nuclei were pelleted by centrifugation at 500 x *g* for 5 minutes. Nuclei were then resuspended in 1mL of NSB and flash frozen into liquid nitrogen in 100 µL aliquots. Nuclei were then stored at -80C until further processing with sci-RNA-seq.

Preparation of sci-RNA-seq2 libraries

Frozen nuclei were thawed over ice and spun down at 500g for 5 minutes. Cells were then permeabilized in permeabilization buffer (NSB + 0.25% Triton-X) for 3 minutes and then spun down. Following another a wash in NSB, two-level sci-RNA-seq libraries prepared as previously described (*25*). Briefly, nuclei were pelleted at 500 x *g* for 5 minutes, and resuspended in 100 µL of NSB. Cell counts were obtained by staining nuclei with 0.4 % trypan blue (Sigma-Aldrich) and counted using a hemocytometer. 5000 nuclei in 2 µL of NSB and 0.25 µL of 10 mM dNTP mix (Thermo Fisher Scientific, cat no. R0193) were then distributed onto a skirted twin.tec 96 well LoBind plate (Fisher Scientific, cat no. 0030129512) after which 1 µL of uniquely indexed oligo-dT (25 µM)(*25*) was added to every well, incubated at 55C for 5 minutes and placed on ice. 1.75 µL of reverse transcription mix (1µL of Superscript IV first-strand buffer, 0.25 µL of 100 mM DTT, 0.25 µL of Superscript IV and 0.25 µL of RNAseOUT recombinant ribonuclease inhibitor) was then added to every well and plates incubated at 55C for 10 minutes and placed on ice. 5 µL of stop solution (40 mM EDTA, 1 mM spermidine and 0.5% BSA) were added to each well to stop the reaction. Wells were pooled using wide bore tips, and nuclei transferred to a flow cytometry tube through a 0.35 µm filter cap and DAPI added to a final concentration of 3 µM. Pooled nuclei were then sorted on a FACS Aria II cell sorter (BD) at 150 cells per well into 96 well LoBind plates containing 5 µL of EB buffer (Qiagen). After sorting, 0.75 µL of second strand mix (0.5 µL of mRNA second strand synthesis buffer and 0.25 µL of mRNA second strand synthesis enzyme, New England Biolabs) were added to each well, second strand synthesis performed at 16C for 150 minutes. Tagmentation was performed by addition of 5.75 µL of tagmentation mix (0.01 µL of a custom TDE1 enzyme in 5.74µL 2x Nextera TD buffer, Illumina) and plates incubated for 5 minutes at 55C. Reaction was terminated by addition of 12 µL of DNA binding buffer (Zymo) and incubated for 5 minutes at room temperature. 36 µL of Ampure XP beads were added to every well, DNA purified using the standard Ampure XP

protocol (Beckman Coulter) eluting with 17 µL of EB buffer and DNA transferred to a new 96 well LoBind plate. For PCR, 2 µL of indexed P5, 2 µL of indexed P7 (*25*) and 20 µL of NEBNext High-Fidelity master mix (New England Biolabs) were added to each well and PCR performed as follows: 75C for 3 minutes, 98C for 30 seconds and 18 cycles of 98C for 10 seconds, 66C for 30 seconds and 72C for 1 minute followed by a final extension at 72C for 5 minutes. After PCR, all wells were pooled, concentrated using a DNA clean and concentrator kit (Zymo) and purified via a 0.8X Ampure XP cleanup. Final library concentrations were determined by Qubit (Invitrogen), libraries visualized using a TapeStation D1000 DNA Screen tape (Agilent) and libraries sequenced on a Nextseq 500 (Illumina) using a high output 75 cycle kit (Read 1: 18 cycles, Read 2: 52 cycles, Index 1: 10 cycles and Index 2: 10 cycles).

Preparation of sci-RNA-seq3 libraries

Frozen nuclei were thawed as before and three-level sci-RNA-seq libraries prepared as described in (*21*). Nuclei were pelleted at 500 x *g* for 5 minutes, washed three times with NSB and a small aliquot of nuclei stained with 0.4 % trypan blue (Sigma-Aldrich) and nuclei counted using a hemocytometer. 80000 nuclei in 22 µL of NSB, 2 µL of 10 mM dNTP mix and were then distributed into a skirted 2 µL of ligation compatible indexed oligo-dT primers were distributed into each well of 96 well LoBind plates, incubated at 55C for 5 minutes and placed on ice. 14 µL of reverse transcription mix (8µL of Superscript IV first-strand buffer, 2 µL of 100 mM DTT, 2 µL of Superscript IV and 2 µL of RNAseOUT recombinant ribonuclease inhibitor) was then added to every well and RT performed on a thermocycler using the following program: 4C for 2 minutes, 10C for 2 minutes, 20C for 2 minutes, 30C for 2 minutes, 40C for 2 minutes, 50 for 2 minutes and 55C for 15 minutes. After RT, 60 µL of nuclei buffer containing BSA (NBB, 10 mM Tris-HCl, pH 7.4, 10 mM NaCl, 3 mM MgCl2 and 1% BSA) were added to each well, nuclei pooled using a wide bore tip, nuclei pelleted by centrifugation at 500 x *g* for 10 minutes and the supernatant removed. A second round of combinatorial indexing was performed by ligation of indexed primers onto the 5' end of RT indexed cDNA. Nuclei were resuspended in NSB and 10 µL added to each well of 96 well LoBind plates after which 8 µL of indexed ligation primers were added to each well along with 22 µL of ligation mix (20 µL of Quick ligase buffer and 2 µL of Quick ligase, New England Biolabs). Ligation was then performed at 25C for 10 minutes. After ligation, 60 µL of NBB were added to each well, nuclei pooled using a wide bore tip, another 40 mL of NBB added to the nuclei and nuclei pelleted by centrifugation at 600 x *g* for 10 minutes and the supernatant removed. Nuclei were then washed once with 5 mL of NBB, resuspended in 4 mL of NBB, multiplets removed by filtering using a 40 µm Flowmi cell strainer (Sigma-Aldrich), nuclei counted and 5000 nuclei were distributed per well into 96 well LoBind plates in a 5 µL volume. Plates containing nuclei were frozen and stored at -80C until further processing. After thawing the frozen plate 5 µL of second strand synthesis mix (3 µL of elution buffer, 1.33 µL mRNA second strand synthesis buffer and 0.66 µL of mRNA second strand synthesis enzyme) were added to each well and incubated at 16C for 3 hours. Tagmentation was performed by addition of 10 µL of tagmentation mix (0.01 µL of a custom TDE1 enzyme in 9.99µL of 2x Nextera TD buffer, Illumina) and plates incubated for 5 minutes at 55C. After tagmentation, 20 µL of DNA binding buffer was added to every well and plates incubated at room temperature for 5 minutes. 40 µL of Ampure XP beads were then added to each well and plates incubated for 5 minutes at room temperature. Upon isolation of beads using a magnetic stand, supernatant was removed and beads were washed twice with 80% ethanol. 10 µL of USER reaction mix (1 µL of 10X USER buffer and 1 µL of USeR enzyme in nuclease-free

water, New England Biolabs) was then added to each well and beads resuspended and incubated at 37C for 15 minutes. After incubation, 7 µL of elution buffer were added to each well and supernatant transferred to a new 96 well LoBind plate after binding beads on a magnetic stand. After incubation at 85C for 10 minutes, libraries were generated with 15 cycles of PCR. Following PCR amplification, sequencing library was purified by first concentrating 1mL of PCR library using a 1x Ampure cleanup and then running the resulting product on a 2% agarose gel containing ethidium bromide. Gel was cut to isolate 2 fragments, hash molecules (220bp - 250bp) and RNA library (250bp - 1000bp). Following gel extraction and an additional 1x Ampure cleanup RNA libraries were sequenced on a NovaSeq 6000 (Illumina) (Read 1: 34 bp, Read 2: 100 bp, Index 1: 10 bp and Index 2: 10 bp) and hash libraries were sequenced on a 75 cycle NextSeq (Read 1: 34 bp, Read 2: 38 bp, Index 1: 10 bp and Index 2: 10 bp).

Preparation of bulk RNA sequencing libraries

Compound treated cells were first trypsinized and harvested as described previously. Cells were then lysed in V-bottom plates using 26 µL of NSB. 2 µL of 25 µM indexed RT primers were added and annealed at 65C for 5 minutes. Subsequently, RT reaction was performed using the SuperScript IV system, with 8µL of 5x SuperScript Buffer, 2µL of SuperScript IV, 2µL 10 mM dNTP mix, 2µL of 100 mM DTT and 2 µL of RNAseOUT recombinant ribonuclease inhibitor per well.. Reaction was performed for 10 minutes at 55C and subsequently stopped via heat inactivation (80C for 10 minutes). Libraries were then pooled and excess RT primer was removed through either two 0.7x SPRI clean-ups or a single 0.7x SPRI cleanup followed by Exo-1 treatment and inactivation. Double stranded DNA was produced through incubation at 16C for 3 hours with second strand synthesis mix containing 0.5 µL of enzyme and 2 µL of second strand reaction buffer in a final volume of 20 µL. Following second strand synthesis, libraries were tagmented with 1 µL of commercial Nextera reagent with 20.5 µL of 2x TD buffer and . Reactions were stopped with 40 µL of Zymo Clean and Concentrate buffer and incubated at room temperature for 5 minutes. Libraries were subsequently purified with a 1x SPRI cleanup and eluted in 16 µL of elution buffer. Sequencing libraries were generated through PCR with 2 µL of index P7 and P5 primers each and 20 µL of 2x NEB Next Master Mix. Finally, libraries were pooled, purified with a 1x SPRI cleanup and quantified. Libraries were sequenced on a Nextseq 500 (Illumina) using a high output 75 cycle kit (Read 1: 18 cycles, Read 2: 52 cycles, Index 1: 10 cycles and Index 2: 10 cycles).

Pre-processing of sequencing data

Sequencing runs were first demultiplexed using bcl2fastq v.2.18. Only barcodes that matched reverse transcription indices within an edit distance of 2 bp were retained. For sci-RNA-seq3 libraries, barcodes which matched both provided reverse transcription indices and ligation indices within an edit distance of 2 bp were retained. Following assignment of indices, polyA tails were trimmed using trim-galore, and reads were mapped to a human transcriptome (hg-38) or human-mouse transcriptome (hg-38 and mm-10) using the STAR aligner. Following alignment, reads were filtered for alignment quality, and duplicates were removed. Reads were considered duplicates if they (1) mapped to the same gene, (2) mapped to the same cell barcode and (3) contained the same unique molecular identifier (UMI). Reads that met the first two criteria, and differed by an edit distance of 1 from a previously observed UMI were also marked as duplicates and discarded. Non-duplicate reads were assigned to genes using bedtools *(40)* to intersect with an annotated gene model. All 3' UTRs in the gene model were extended by 100 bp

to account for the possibility that some gene 3' UTR annotations may be too short, causing genic reads to improperly be annotated as intergenic. Cell barcodes were considered to correspond to a bona fide cell if the number of unique reads associated with the barcode was greater than an interactively defined threshold on a knee plot. Reads from cells that passed this UMI count threshold were first aggregated into a sparse matrix format and then loaded and saved as a CDS object for analysis with Monocle 3.

Assigning sample labels from hash reads

Demultiplexed reads that matched combinatorial indexing barcodes were examined to identify hash reads. Reads were considered hash reads when they met two criteria: (1) the first 10 bp of read 2 matched a hash barcode in the experiment within an edit distance of two and (2) contained a polyA track between base pairs 12 to 16 of read 2. These reads were then deduplicated by cell barcode and collapsed by UMIs to create a vector h_i of hash oligo UMI counts for each nucleus i in the experiment.

To assign each nucleus i to the culture well from which it came, we test whether its sci-RNA-seq library is enriched for a particular hash barcode. We compare a nucleus's hash UMIs against a 'background distribution', which under ideal circumstances, would be the uniform distribution. In practice, minor variation in concentrations of hash oligos added to each well of liberated nuclei may necessitate empirically estimating the background. To do so, we simply average the relative hash UMIs from cell indices for which fewer than < mRNA UMIs were collected, reasoning that these reflect library contributions from RT well supernatant, debris fragments, etc. We then compare the hash UMIs h_i for nucleus i to this background by a chi-squared test. After correcting the resulting p values for multiple testing by Benjamini-Hochberg, we reject the null hypothesis that h_i originates from the background distribution at specified FDR (5% FDR was used in this study). Those nuclei with hash counts deemed different than background are then evaluated for enrichment for a single hash sequence. Enrichment ratios were calculated as the UMI count ratio of the most abundant vs. the second most abundant hash oligo. Specifically, if the UMI count for the most abundant hash in nucleus i is α-fold higher than the second most abundant, i is marked as a singleton. α was determined on a per-experiment basis by examining the distribution of these ratios and choosing a value that separated unlabeled cells and singularly labeled cells. Cells that fell below α-fold enrichment of a unique hash oligo were flagged as a multiplet or debris and discarded.

Dose-response analysis

Dose-response analysis was conducted in R using the drc package (*41*) by fitting a four-parameter log-logistic model for each drug to the number of cells recovered in the single-cell RNA-seq data at each dose. Cells that survived doublet analysis and QC were grouped by their culture well of origin and counted. These counts were then adjusted to account for variation in recovery as a function of cell type and culture plate as follows. The vector x of cell counts across wells were fit with the model

$$\ln(x_i) = \beta_0 + \beta_{t_1} t_1 + \ldots + \beta_{t_m} t_m + \beta_{w_1} w_1 + \ldots + \beta_{w_n} w_n$$

Where t_i and w_j are binary indicator variables encoding the cell type and culture plate, respectively. The adjusted cell counts for a given well J from culture plate p of cells of type k are then computed as

$$\tilde{x}_j = x_j - \exp(\beta_{t_k} t_k + \beta_{w_p} w_p)$$

Next, adjusted per-well cell counts were grouped by type and drug and passed as input to the drm() function of the drc package with a model formula 'cell_count ~ log_dose` and the LL.4() model family function. This procedure fits the model:

$$f(x; b, c, d, e) = c + \frac{d - c}{1 + \exp(b(\ln x - \ln e))}$$

In the above model, the parameters c and d correspond to the lower and upper asymptotic limits of the response, respectively. The steepness of the response curve is reflected in b, and $e$is a parameter that encodes the half-maximal 'effective dose' (ED50).

The dose response curves enable cells to be annotated according to the impact of their culture conditions on viability. Each cell is assigned a 'viability score' which is simply the expected fraction of vehicle cells remaining after exposure to a given dose of a compound. These cell counts are generated via the predict() function of the drc package and then normalized relative to the corresponding vehicle control.

Dimensionality reduction and trajectory analysis

Gene expression profiles were visualized with Monocle 3, which uses UMAP to project them into a two or three dimensional space. Briefly, Monocle 3 first calculates size factors for every cell. Size factors were calculated as the log UMI counts observed in a single cell divided by the geometric mean of log UMI counts from all measured cells. After scaling each nucleus' UMI counts by its library size factor, Monocle3 adds a pseudocount of 1, and log transforms the counts. Next, these log-transformed profiles are projected onto the top 25 principal components. These PCA coordinates were transformed by Monocle 3 (using an approach similar to the removeBatchEffect() function in the limma package (*42*)) according to the model '~ log(UMIs) + replicate' (**Figure 3**) or '~ log(UMIs) + viability + proliferation index + replicate' . Adjusted PCA coordinates for each cell are used to initialize UMAP. Unless otherwise noted, UMAP was run with the following parameters: 50 nearest neighbors, min_dist = 0.1, inter-cell distance assessed by cosine similarity. UMAP projection of cells after dual HDAC inhibition and acetyl-COA precursor supplementation or acetyl-CoA generating enzyme inhibition was performed as described with the exception that PCA initialization was performed on the top 1000 most overdispersed genes. Louvain community detection was then performed on this UMAP space using the python package 'louvain'. Trajectory reconstruction was then performed as described in (*21*).

To determine whether cells exposed to a particular compound/dose combination displayed an enrichment along UMAP space we created contingency tables of the number of compound or vehicle treated cells within and outside clusters and used the stats R package implementation of Fisher's exact test to test for enrichment. For visualization of drug enrichment in **Figure 3B**,

cells opacity was added to cells under the minimum compound/dose that passed meet an enrichment cutoff of FDR < 1% and a log2 of the odds ratio > 2.5. Cells that passed these filters were used to generate the heatmap of the fraction of enriched cells by cluster in **Figure S6**.

Estimation of Proliferation Index

To obtain an estimate of proliferation index for a single cell, size factor normalized expression of cell cycle marker genes (from Table S5 in (*43*)) were summed for each cell and logged. Scores were calculated in this way for both G1S and G2M. "Proliferation Index" refers to overall proliferative state of a cell and is calculated as the logged sum of the aggregated G1S and G2M gene expression.

Differential expression analysis

To test whether a gene is differentially expressed by a cell line in a dose-dependent manner when exposed to a compound, we fit its (library size-factor adjusted) UMI count recorded from each nucleus with a generalized linear model:

$$\ln(y_i) = \beta_0 + \beta_d d$$

Where y_i is a quasipoisson-valued random variable, d is the log-transformed dose of the compound being evaluated. We fit these models with Monocle 3, which uses the speedglm package. To fit the regression model for each drug's effect on each gene, we first identify the subset of cells that are relevant for the model. To determine the effects on gene G in cells of type C when treated with drug D, we include all cells of type C that were treated with any dose of D. To these, we add cells of type C that were treated with the vehicle control. We then fit a model defined above relating the expression level of G across all of these cells. Genes are deemed to be dose-dependent differentially expressed genes (DEGs) if their fitted models include a term β_d that is significantly different from zero as assessed by a Wald test (Benjamini-Hochberg adjusted $p < 0.05$). P values for β_d terms are pooled across all compounds and all genes prior to correction for multiple testing.

To assess a gene for differential expression as a function of 'pseudodose' ψ in the consensus HDAC inhibition trajectory, we fit a model

$$\ln(y_i) = \beta_{\tilde{\psi}}\tilde{\psi} + \beta_c c + \beta_{c\tilde{\psi}}\tilde{\psi}c + \beta_{d_1} d_1 + \ldots + \beta_{d_k} d_k$$

Where y_i is a quasipoisson variable capturing the gene's UMI counts, $\tilde{\psi}$ encodes the pseudodose values smoothed via a natural spline, $c$is a factor encoding the cell type, and $\beta_{c\tilde{\psi}}$ captures the interaction between cell type and pseudodose. The term $\beta_{d_j} d_j$ encodes the (log) dose dependent effects of compound j.

Pairwise correlation of screened compounds

To identify compounds that result in similar dose-dependent changes to cellular transcriptomes we calculated the Pearson correlation between every pairwise set of compounds. We created a

gene by compound matrix for the union of dose-dependent genes across all compounds where each entry is the beta coefficient for the dose dependence term β_d and ten calculated the Pearson correlation for every drug pair using the cor.test() function in the R stats package specifying to use complete observations. The resulting correlation matrix was then hierarchically clustered using the pheatmap package in R. The significance of every pairwise correlation was determined using the corr.test() function from the psych package in R specifying Benjamini-Hochberg as the method for adjusting for multiple hypothesis testing.

Geneset enrichment analysis

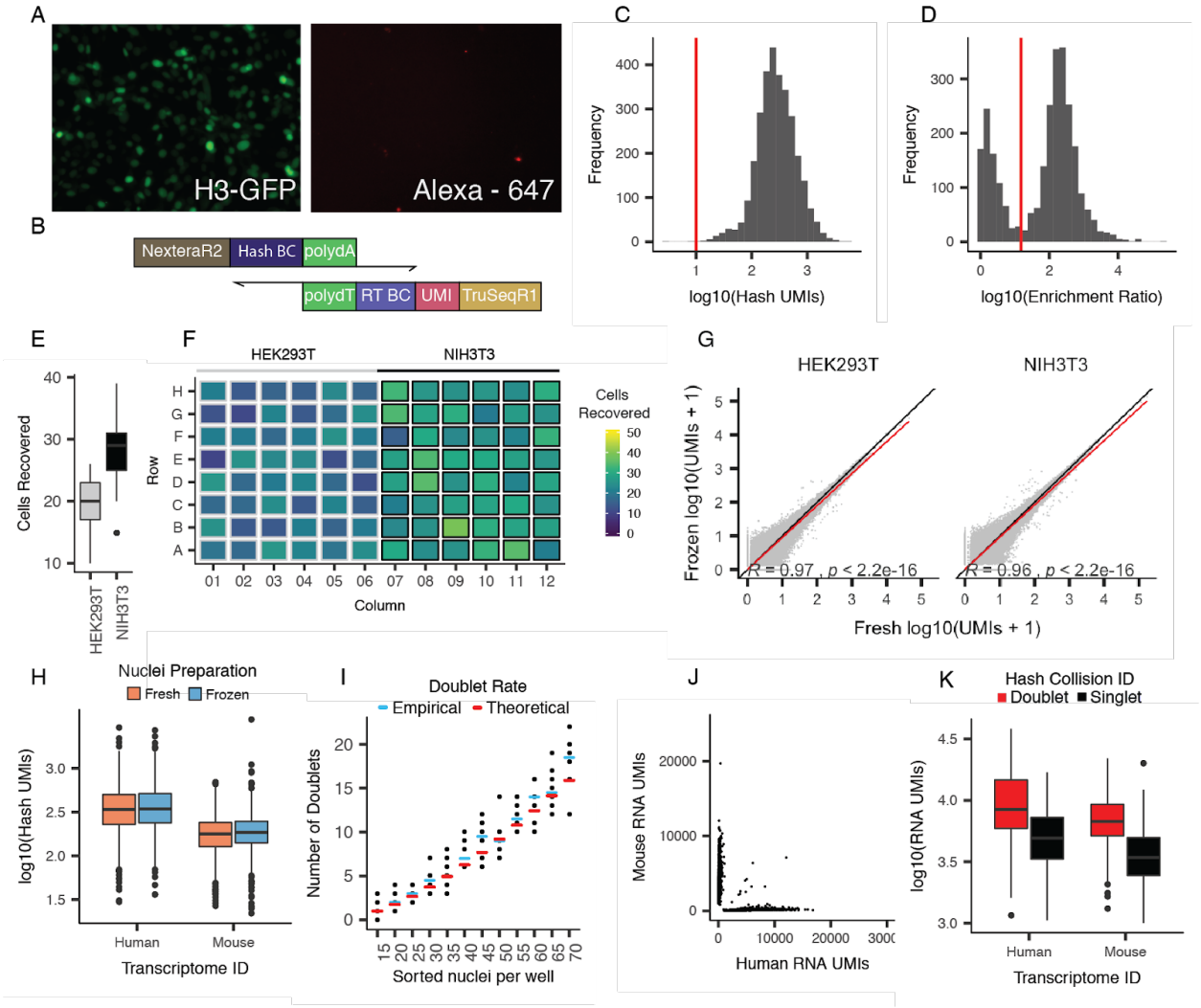
After fitting a generalized linear model, genes that had significant coefficients (5% FDR threshold) were used for gene set enrichment analysis with the R package piano (*44*). Briefly, gene sets were ranked according to the set-wide average Wald test statistic corresponding to the generalized linear model term being evaluated with piano's runGSA() function. Genes were randomized across sets to establish a null distribution for each set's rank. After 10000 permutations, runGSA() computed p values using the 'mixed' directional enrichment policy.. The top gene sets, corresponding to those with the largest magnitude enrichment statistic, were chosen for visualization.

Alignment of HDAC inhibitior treated cells

To organize cells treated with HDAC inhibitors into a trajectory cells were sampled to equalize the number of cells represented between the three cell lines or between treatments at 24 and 72 hrs. Next, PCA coordinates were computed jointly, and then aligned using the mnnCorrect function from the package scran (*32*). These adjusted coordinates were used to initialize UMAP in Monocle 3. We then fit a principal graph to the data via lean_graph(). To define the origin of the trajectory, we mapped each cell to its nearest principal graph node, and then selected all principal graph nodes for which a majority of mapped cells were treated with vehicle. All other cells' pseudodoses ψ was measured as the geodesic distance between their nearest principal graph node to an origin node.

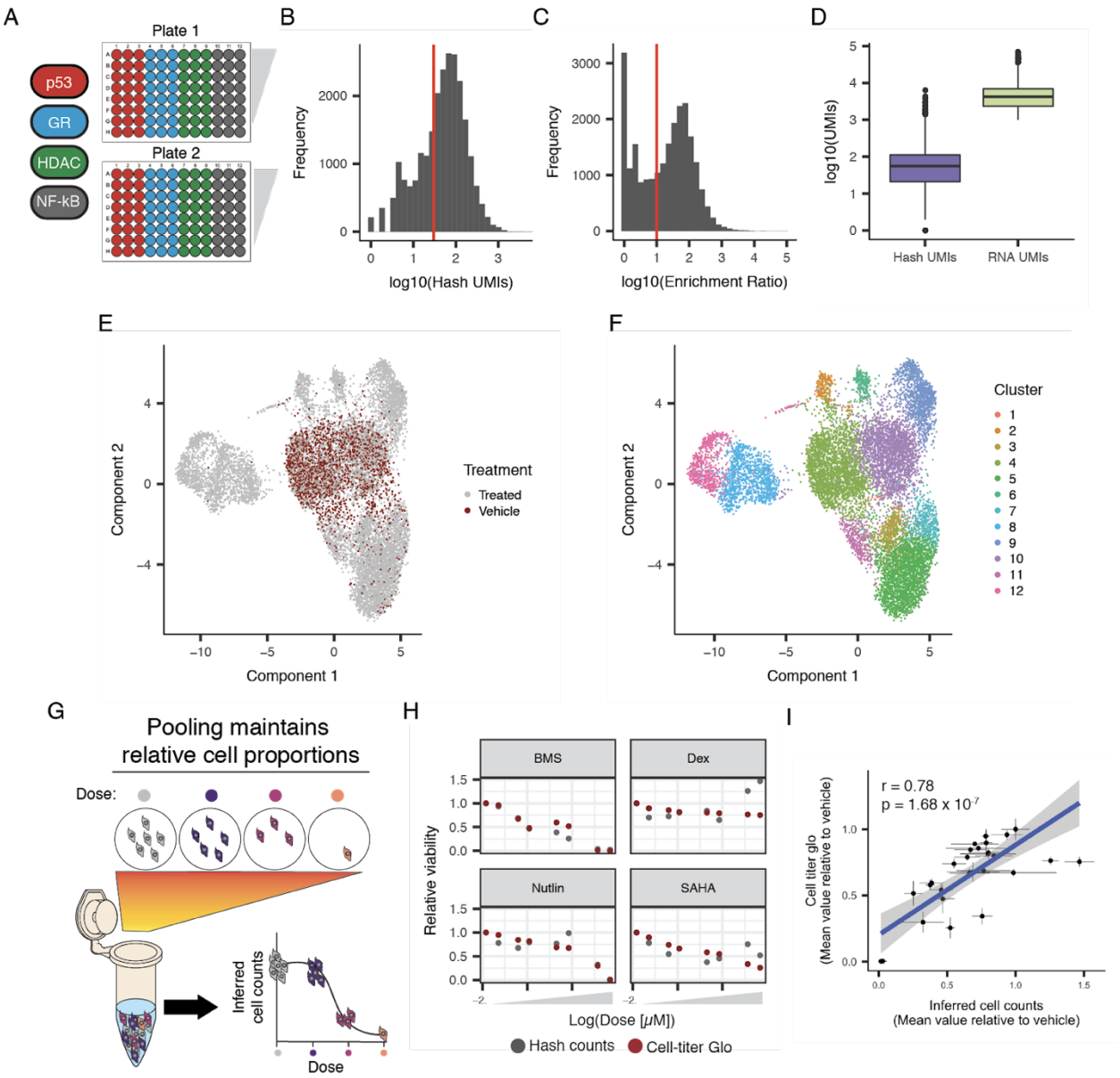
To quantify the potency of each HDAC inhibitor, we first grouped all cells from each replicate according to treatment and dose, and then computed the mean pseudodose for each cell. We then fit mean pseudodose values as a function of compound concentration using the drc package (*41*). We used a four-parameter log-logistic model, with the maximal response fixed at the highest pseudodose value achieved across all compounds and doses. We then take the model parameter e as described in the 'dose response analysis' section above as the transcriptional EC50 (TC50) for each compound.

Supplementary Figures



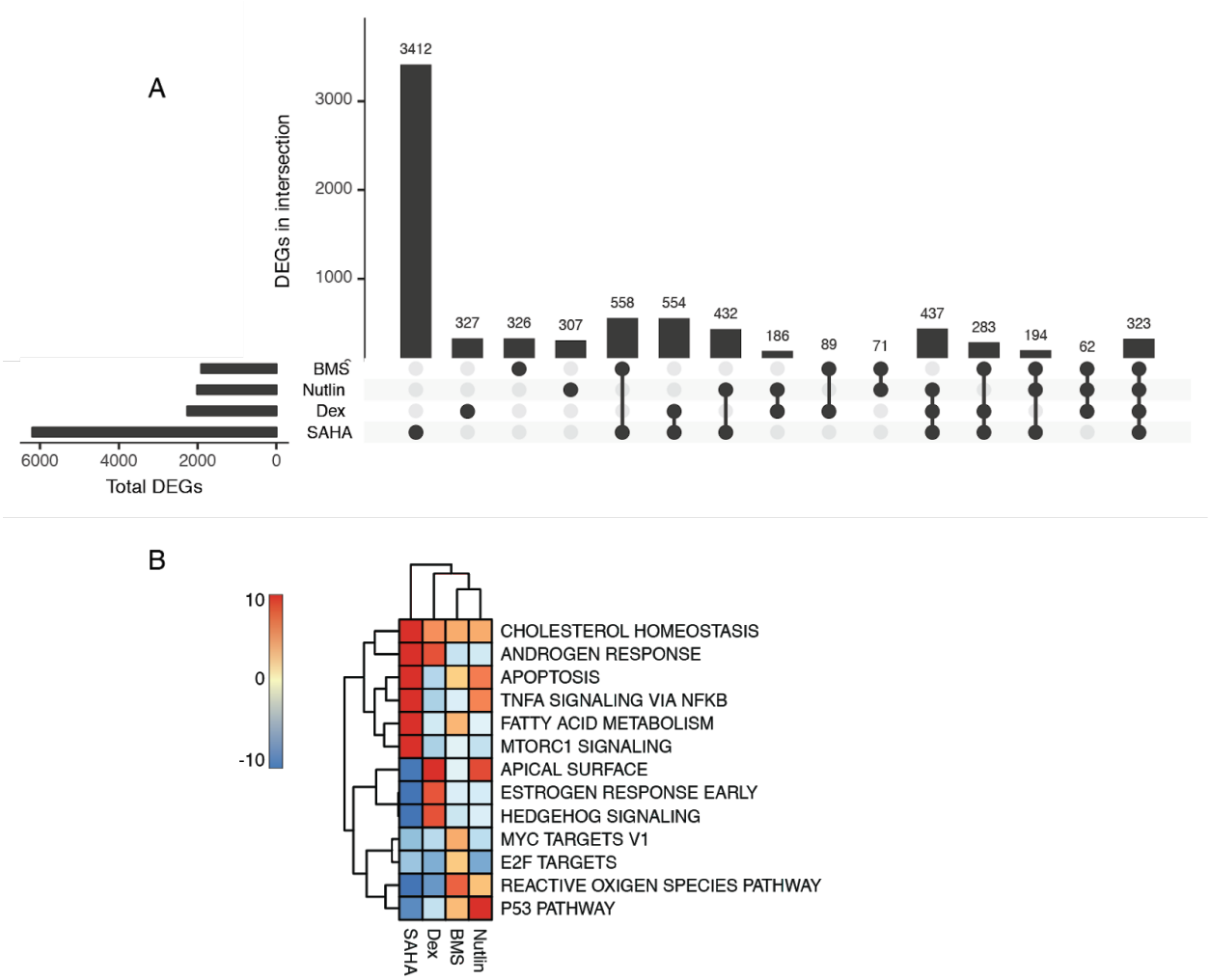
Supplementary Figure 1. **Hashing with short, polyadenylated single-stranded oligonucleotides enables stable, low-cost labeling of nuclei for sci-RNA-seq and subsequent doublet detection. A)** Fluorescent microscopy images demonstrating lack of Alexa 647-conjugated oligo staining (right) of unpermabilized H3-GFP^{+} NIH3T3 cells (left). **B)** Design of polyadenylated hash oligos (top) and indexed primer used for reverse transcription (bottom). **C)** Number of hash UMIs detected per cell. Cells with fewer than 10 hash UMIs (red line) were excluded from further analysis. **D)** Distribution of enrichment ratios for cells. Enrichment ratios were calculated as the UMI count ratio of the most abundant vs. the second most abundant hash oligo. An enrichment ratio cutoff of 15 (red line) was used to distinguish doublets vs. singlets. **E)** Boxplot of the number of cells recovered per well for each cell line. **F)** Layout of culture plate wells with color indicating number of cells recovered and outline indicating cell line. Note that although more NIH3T3 cells were recovered per well, similar numbers of cells were recovered across wells of each cell type. **G)** Log-scale per-gene aggregated, size-factor normalized UMI counts recovered from sci-RNA-seq on fresh vs. frozen preparations. Size factors are calculated as the log counts observed in a single cell divided by the geometric mean of log counts from all measured cells. Black line indicates y = x. Red line is the fit with Pearson correlation shown. **H)**

Log-scale boxplot of number of hash UMIs recovered from sci-RNA-seq of HEK293T (human) or NIH3T3 (mouse cells) from fresh vs. frozen preparations. **I)** Theoretical (red bars) vs. observed (black dots for individual wells and blue bars for means) doublet rate as a function of the number of nuclei sorted into the final plate during sci-RNA-seq. **J)** Barnyard plot from **Figure 1E** after removal of doublets detected by hashing. **K)** Log-scale boxplot of number of RNA UMIs in singlet vs. doublet cells, as called based on the purity of hash UMIs. Of note, these are 'within species' doublets, *i.e.* human-human or mouse-mouse, which are not readily detected by conventional barnyard experiments.

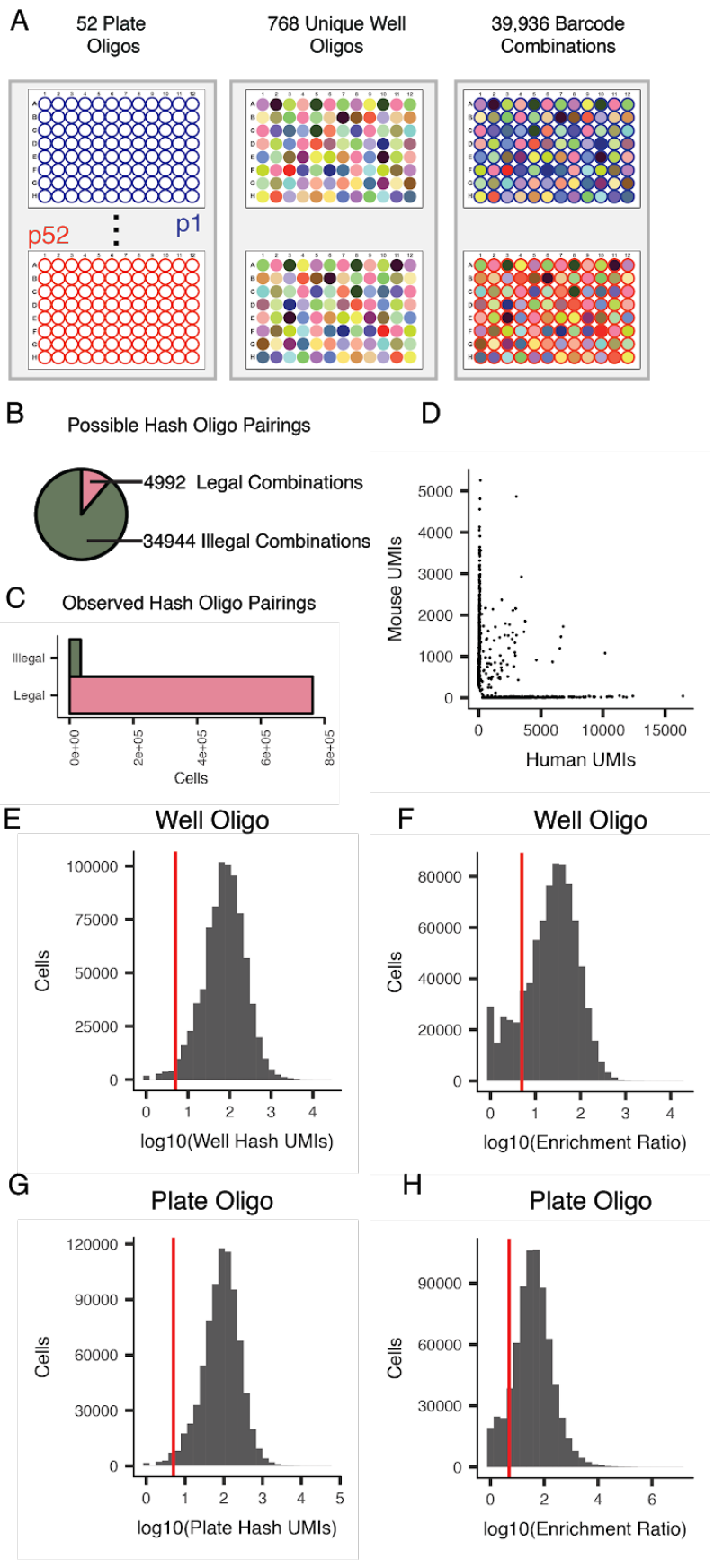


Supplementary Figure 2. sci-Plex distinguishes transcriptional responses of A549 cells to four small molecules and recovers dose-response estimates similar to established assays. **A)** Experimental layout of A549 cells in 96 well plates. Cells were treated for 24 hours in two 96 well plates using 7 doses (or vehicle) arrayed along each column. **B)** Cells that contained more than 30 hash oligo UMIs and **C)** had an enrichment ratio of greater than 10 were retained. **D)** Retained cells had a median hash UMI count of 78 and median RNA UMI count of 4,681. **E)** UMAP embedding of chemically perturbed A549 cells, equivalent to **Figure 2B** but with cells colored by whether they were treated with vehicle or one of the four small molecules. **F)** UMAP embedding of chemically perturbed A549 cells, equivalent to **Figure 2B** but with cells colored by cluster as defined using the density peak algorithm in Monocle 3. **G)** Cartoon depicting how pooling of barcoded nuclei preserves relative cell counts. **H**) Viability estimates from counting the proportion of recovered hashed nuclei (grey) vs. CellTiter-Glo (red, n = 6). **I**) Scatter plot of

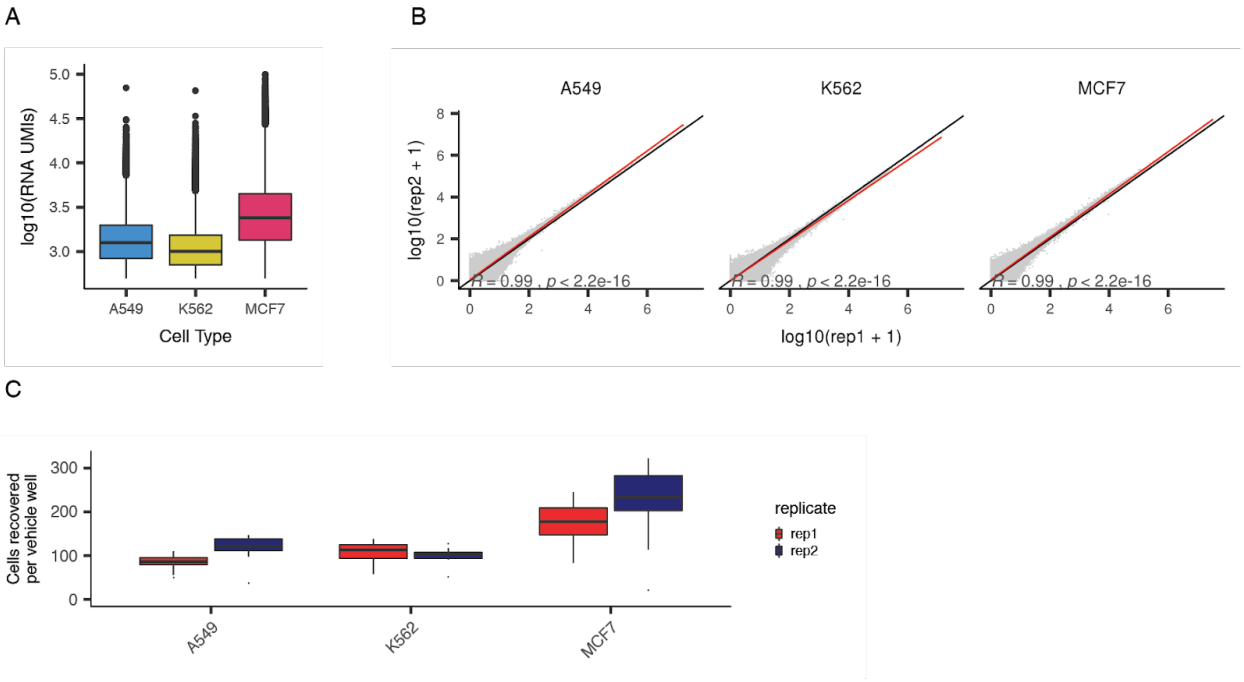
inferred cell counts (x-axis) and CellTiter-Glo viability estimates (y-axis) across all treatments and doses tested (Pearson correlation and chi square test).



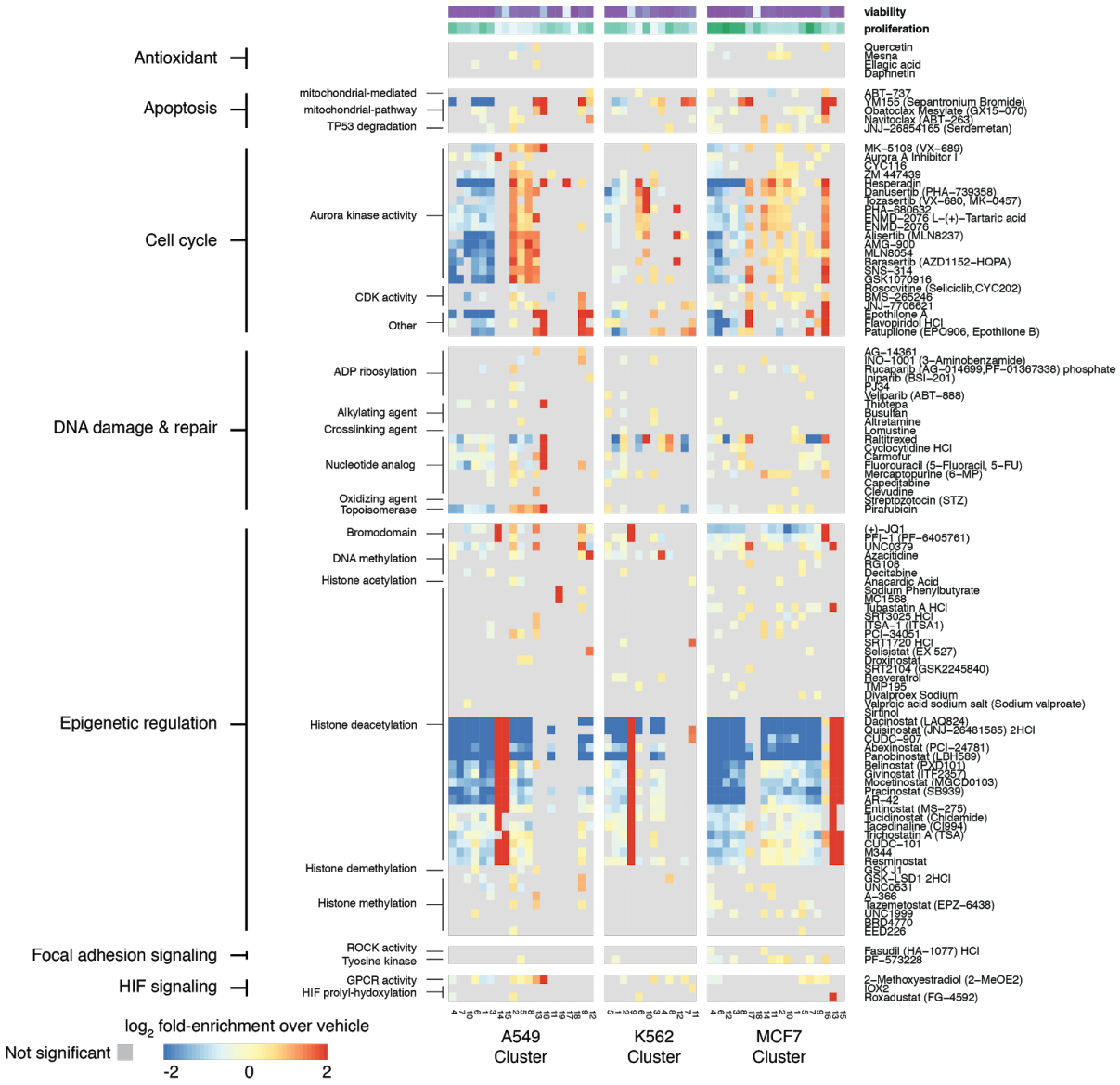
Supplementary Figure 3. Dose-dependent differentially expressed genes (DEG) recover expected transcriptional modules. A) Upset plot displaying the intersections of dose-dependent DEGs between treatments (vertical bars) as well as the total number of dose-dependent DEGs per treatment (horizontal bars). A gene is defined as a dose-dependent DEG if the quasi-poisson regression model relating its expression in a given cell to the dose of drug that cell received shows a significant dose effect (Wald test) after Benjamini-Hochberg correction (FDR < 0.05). See **Methods** for full details on regression modeling. The four leftmost vertical bars correspond to drug-specific dose-dependent DEGs, while the rightmost vertical bar corresponds to dose-dependent DEGs shared by all four drugs. **B)** Gene set analysis (GSA) performed with dose-dependent DEGs using the runGSA() function from the piano package and the Hallmarks gene set from MSigDB (*45*). Heatmap color indicates the value of the directional GSA enrichment statistic with values that were capped at either -10 or +10 for visualization.


A
52 Plate Oligos
768 Unique Well Oligos
39,936 Barcode Combinations
p52
p1
B
Possible Hash Oligo Pairings
4992 Legal Combinations
34944 Illegal Combinations
C
Observed Hash Oligo Pairings
Illegal
Legal
0e+00
2e+05
4e+05
6e+05
8e+05
Cells
D
Mouse UMIs
Human UMIs
E
Well Oligo
Cells
log10(Well Hash UMIs)
F
Well Oligo
Cells
log10(Enrichment Ratio)
G
Plate Oligo
Cells
log10(Plate Hash UMIs)
H
Plate Oligo
Cells
log10(Enrichment Ratio)

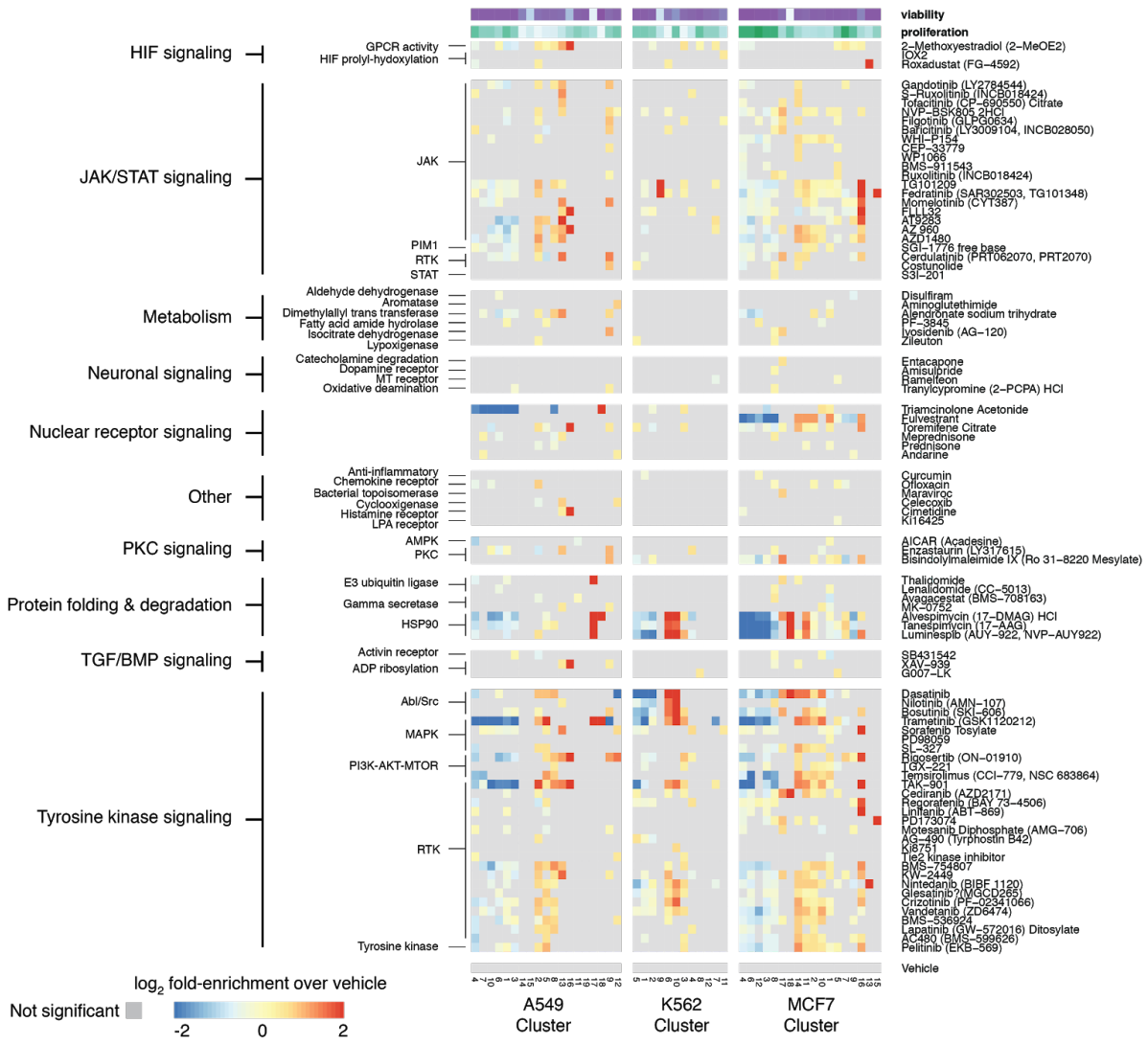
Supplementary Figure 4. Hash-based cell labeling in large-scale sci-Plex experiment. **A)** Hashing design for sci-Plex with 188 compounds. The experiment used 52 x 96-well plates where each well was marked by a combination of two oligos, one specific to a single 96-well culture plate and another specific to a well within that culture plate. **B)** Although this could theoretically be implemented with just 96 well hash oligos, we instead used 768, which meant that out of the 39,936 possible pairings of plate and well hash oligos, only a minority (12.5%) of combinations were expected ('legal'), while most were unexpected ('illegal') **C)** Observed pairings of plate and well hash oligos were strongly enriched for 'legal' combinations. **D)** Scatter plot of HEK293T and NIH3T3 cells seeded in a single RT well of the large-scale sci-Plex experiment. **E-H)** Hash UMI (panels E & G) and enrichment ratio (panels F & H) cutoffs used for well hash oligos (panels E & F) and plate hash oligos (panels G & H). Enrichment ratio cutoffs corresponds to greater than 5-fold enrichment. Hash UMI cutoffs correspond to > 5.



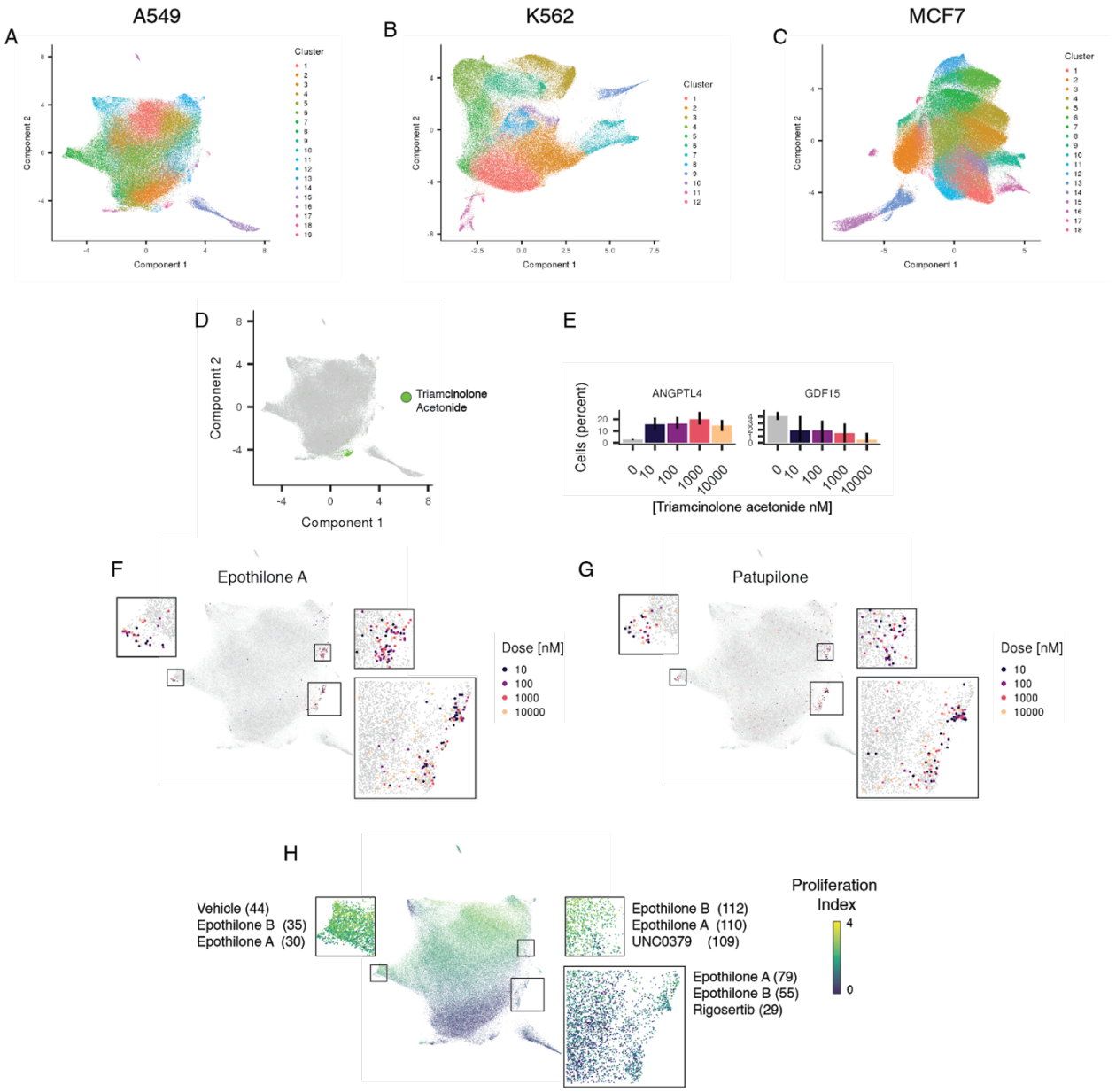
Supplementary Figure 5. Quality control metrics for large-scale sci-Plex experiment. A) Log-scale boxplot of number of RNA UMIs for cells that passed hash and RNA UMI cutoff filters for each of three cell lines. **B)** Correlation of size factor-normalized counts for genes between replicates for each of the three cell lines. Black line indicates y = x. Red line is the fit with Pearson correlation shown. **C)** Boxplots showing the number of vehicle cells recovered from each of 8 vehicle control wells within each replicate for A549, K562 and MCF7 cells.



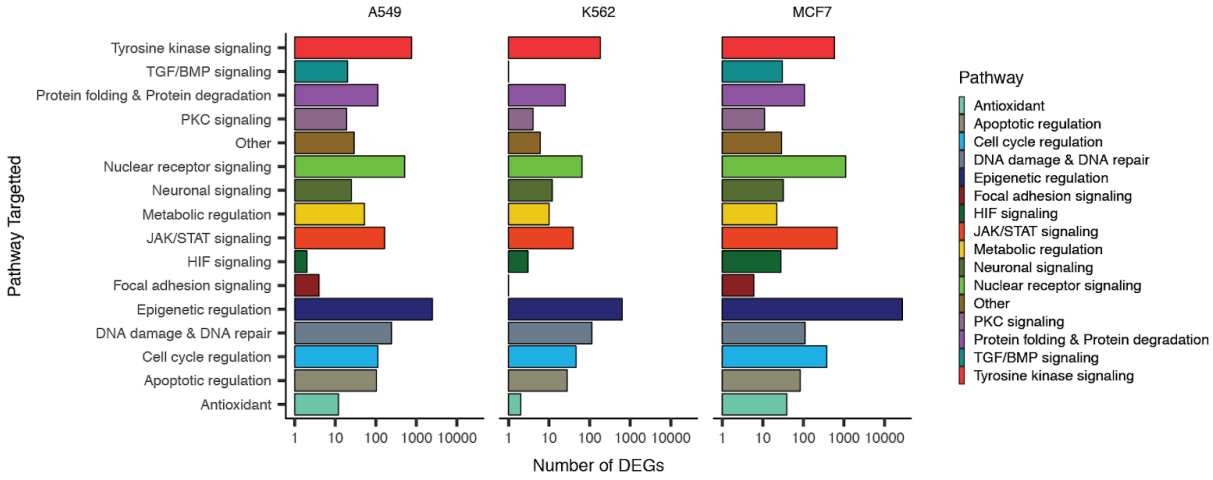
Supplementary Figure 6. Exposing cells to compounds alters their distribution across cell clusters. Heatmap showing the log-transformed ratio of cells treated with a particular drug compared to vehicle control cells in each Louvain community. Columns correspond to clusters in PCA space (see **Fig S7A-C**) and rows correspond to compounds, annotated by pathway and target. A gray entry denotes a compound that is not significantly enriched or depleted relative to vehicle in the corresponding cluster (Fisher's exact test, FDR < 1%).



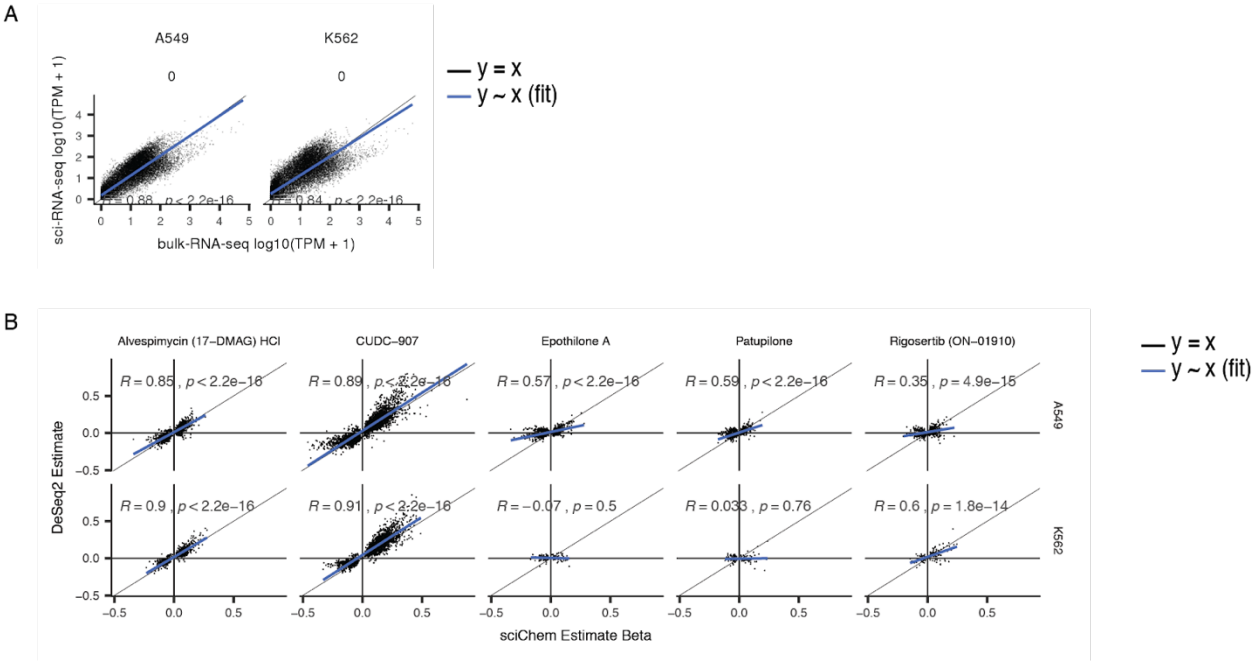
Supplementary Figure 6 (continued). Exposing cells to compounds alters their distribution across cell clusters. Heatmap showing the log-transformed ratio of cells treated with a particular drug compared to vehicle control cells in each Louvain community. Columns correspond to clusters in PCA space (see Fig S7A-C) and rows correspond to compounds, annotated by pathway and target. A gray entry denotes a compound that is not significantly enriched or depleted relative to vehicle in the corresponding cluster (Fisher's exact test, FDR < 1%).



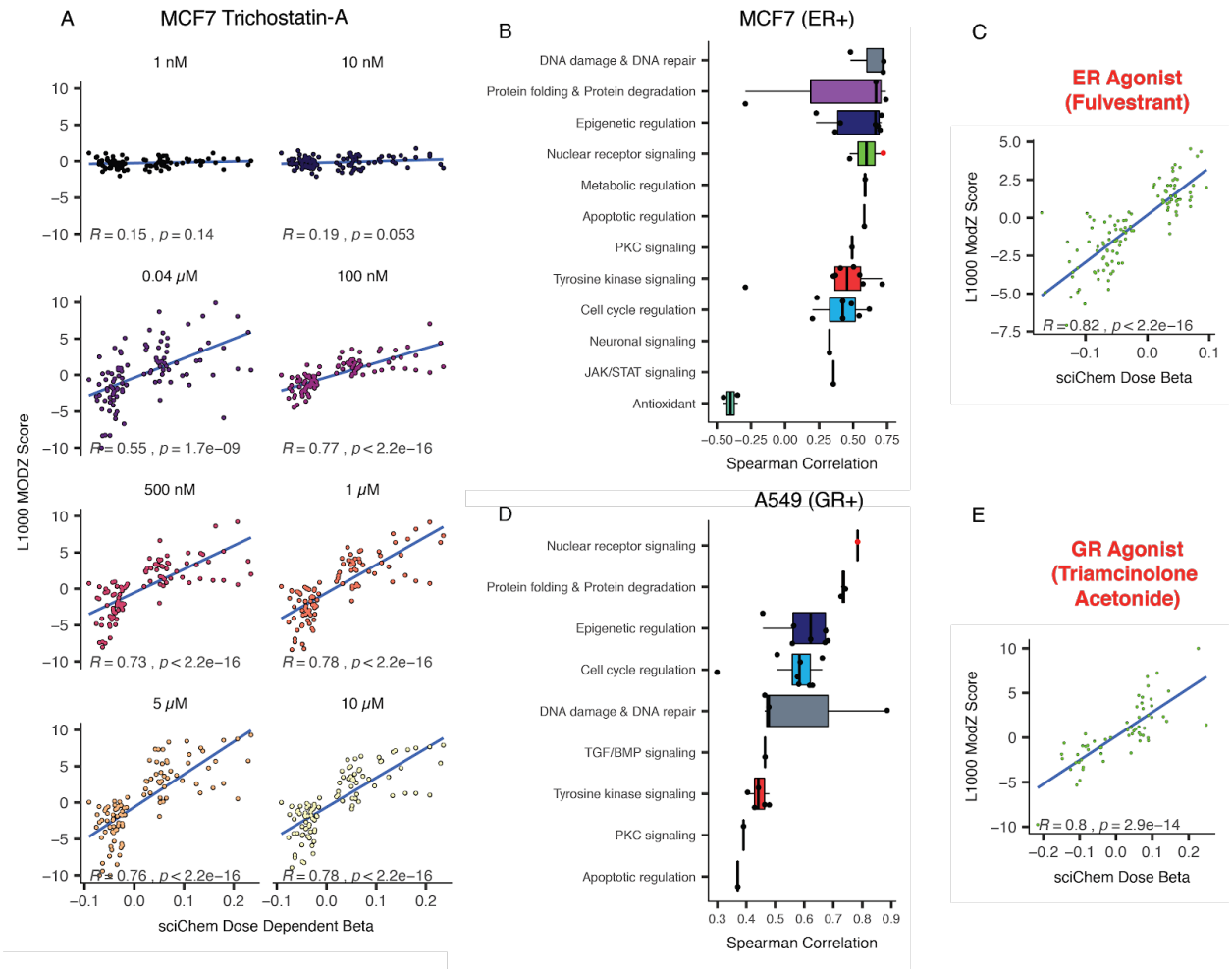
Supplementary Figure 7. sci-Plex identifies pathway-specific enrichment of compounds across UMAP clusters. A-C) UMAP embedding from **Figure 3B** colored by cells' assignment to Louvain communities across PCA space for A549 (panel A), K562 (panel B) and MCF7 (panel C) cells. **D)** UMAP embedding of A549 cells from **Figure 3B**. Cells treated with the glucocorticoid receptor (GR) agonist triamcinolone acetonide are highlighted in green while all other cells are colored grey. These cells comprise the vast majority (95%) of the cells in cluster 18 from panel A. **E**) Percent of A549 cells expressing the GR target genes *ANGPTL4* and *GDF15*, as a function of increasing doses of the synthetic GR agonist triamcinolone acetonide. **F-H)** UMAP embedding of A549 cells colored by cells treated with varying doses of epothilone A **(F)**, epothilone B **(G),** or colored by proliferation index **(H)**. Insets display magnified views of distinct foci induced upon treatment. The treatments with the highest number of cells in each bounding box are indicated in panel H with the number of cells in parentheses.



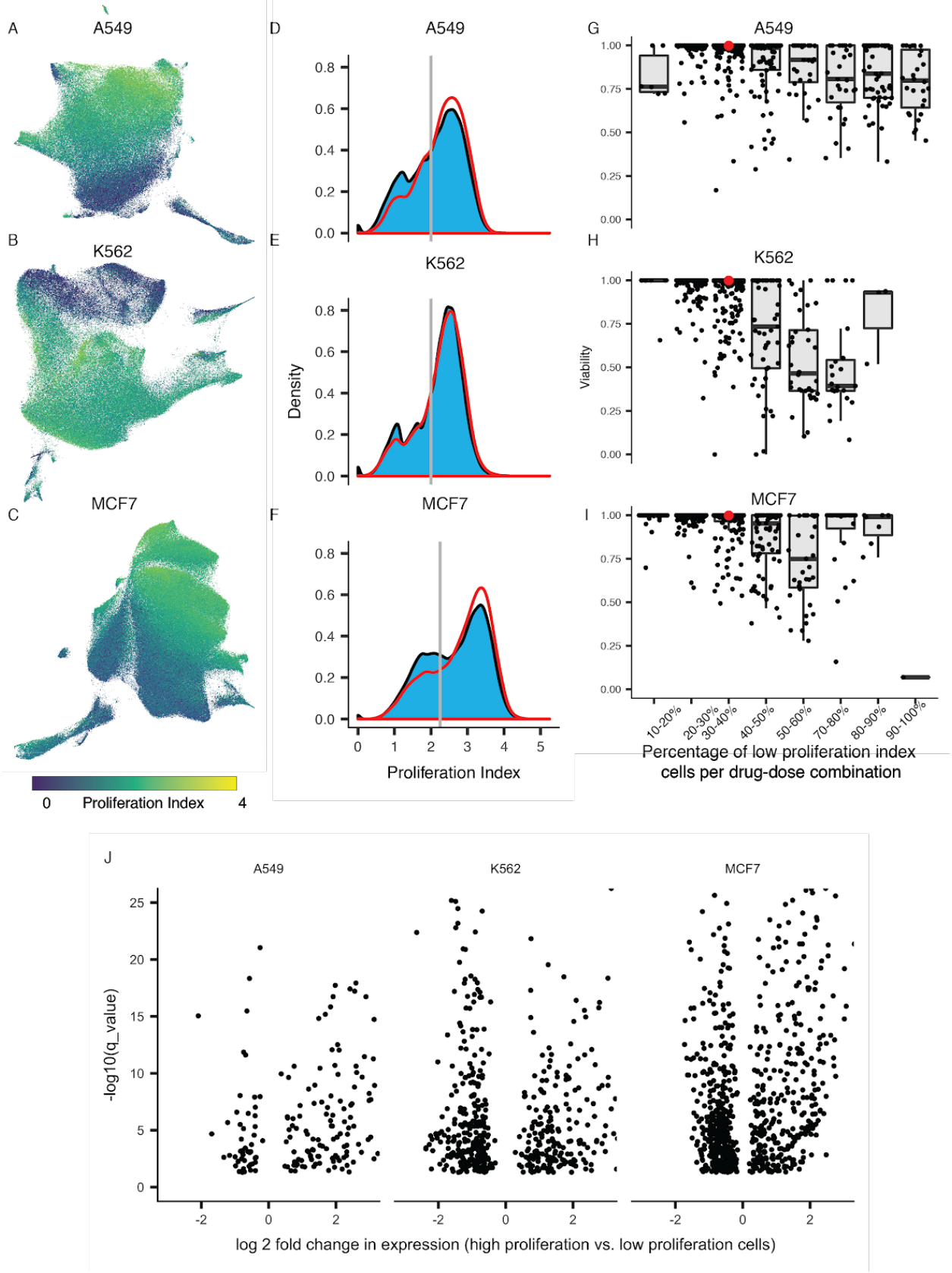
Supplementary Figure 8. Number of dose-dependent differentially expressed genes detected per compound category. Significant dose-dependent differentially expressed genes (FDR < 0.05) are grouped by cell line and colored by targeted pathway.



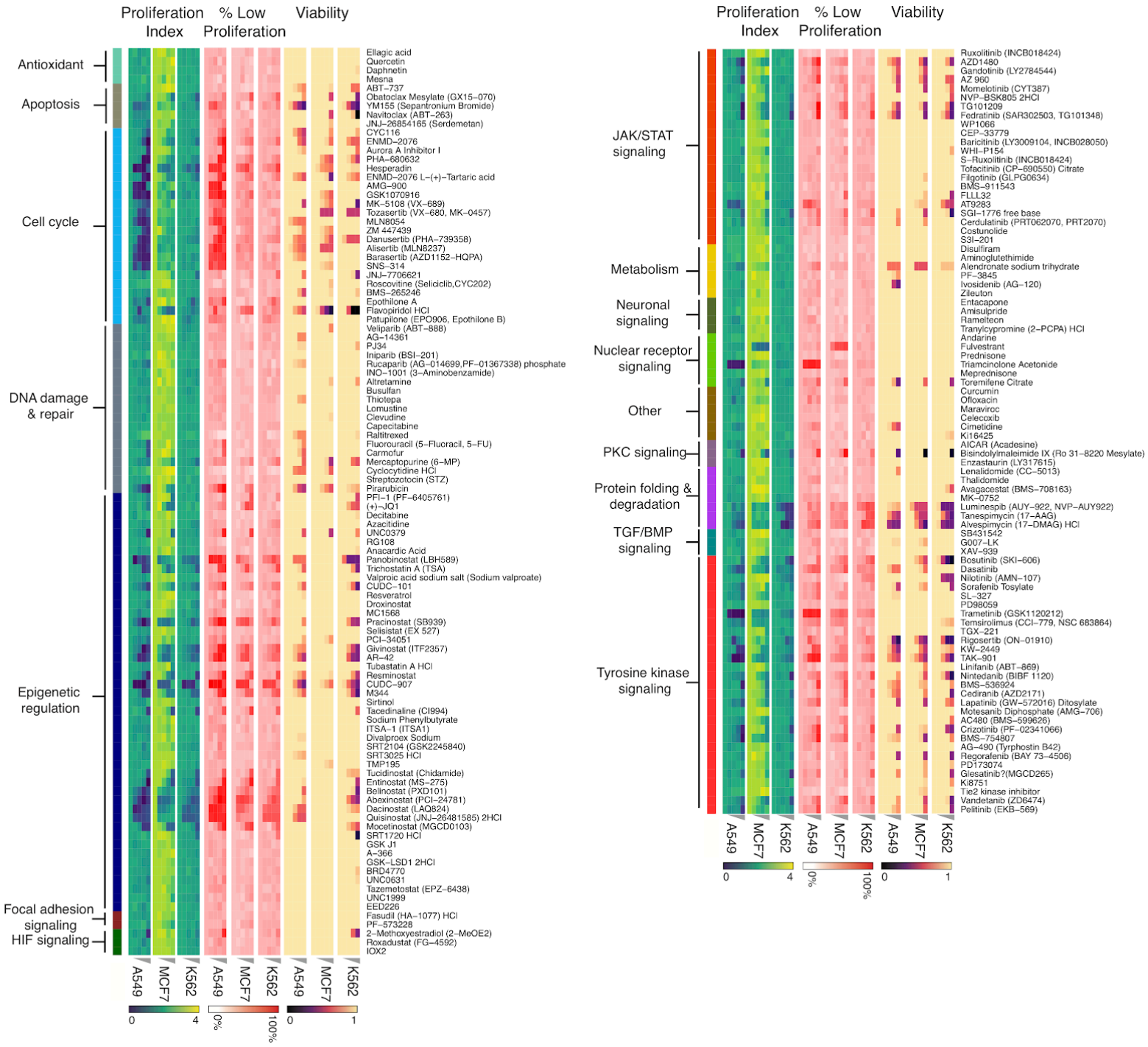
Supplementary Figure 9. Correlation of "pseudobulk" sci-Plex with bulk-RNA-seq. A) Log_{10} transcripts per million (TPM) for protein-coding genes measured by bulk RNA-seq (x-axes) vs. size factor-normalized, aggregated single cell profiles for vehicle treated cells from sci-Plex (y-axis). Results are shown for both A549 and K562 cells. Black line indicates the line y = x, while the blue line shows the linear fit with Pearson correlation shown. **B)** Scatter plots, for selected compounds, comparing statistically significant estimates derived from linear models fit to single cell data (x-axes) vs. estimates derived from bulk RNA-seq using DESeq2 (y-axes). Black line indicates y = x. Blue line is the fit with Pearson correlation shown.



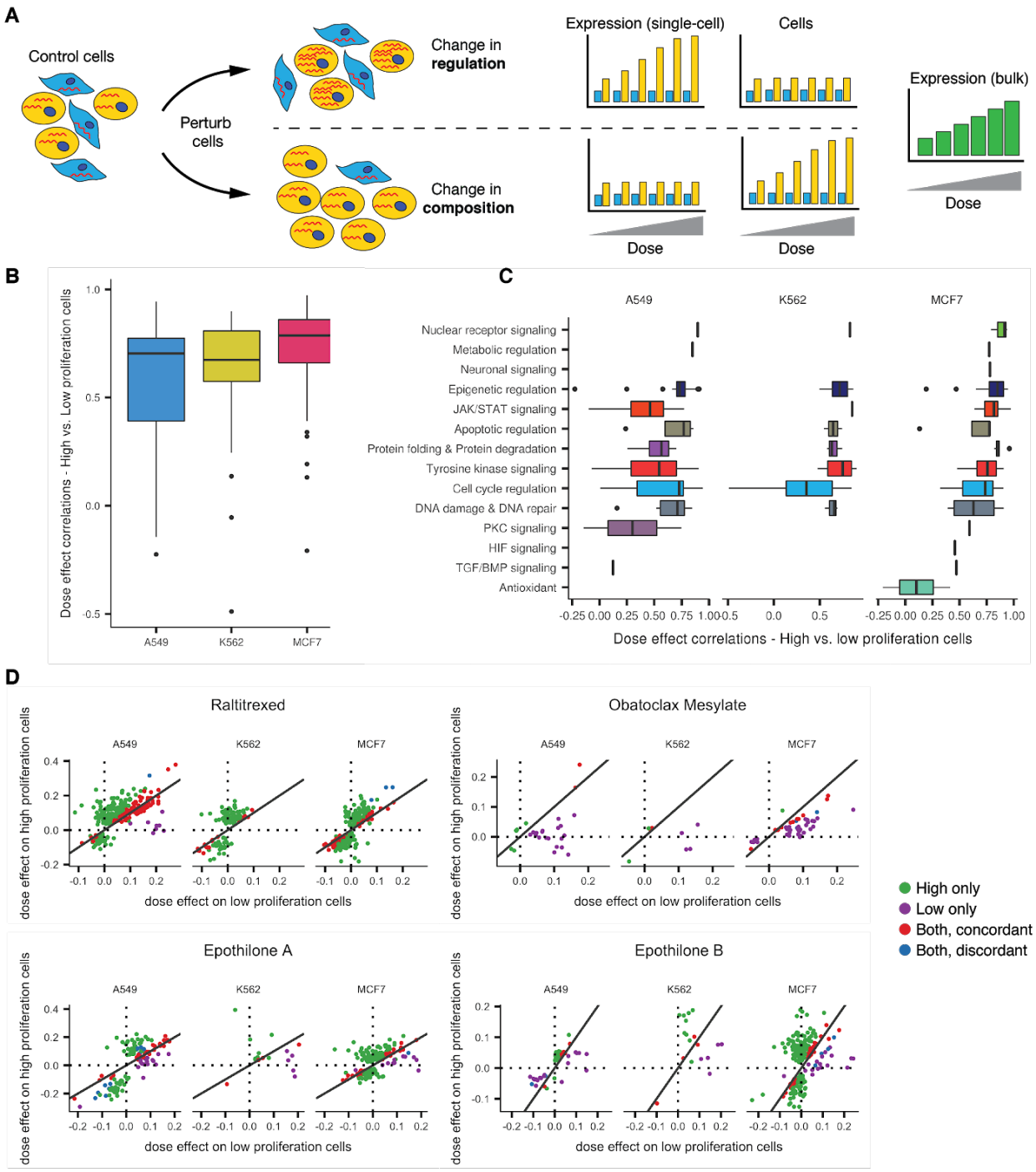
Supplementary Figure 10. Moderated Z scores from the L1000 assay correlate with dose-dependent betas from sci-Plex. A) For a selected compound-cell line combination (trichostatin A in MCF7 cells), we plot moderated Z scores from the L1000 assay with treatment for 24 hrs at each of eight doses (y-axes) (*11*) vs. dose-dependent betas from sci-Plex data (x-axes). All genes that are part of the L1000 assay and significant for dose-dependent effects with sci-Plex (p-value < 0.01) are shown. Line is the fit with Spearman correlation shown. **B)** Boxplot of Spearman correlations between significant sci-Plex computed dose-dependent betas and L1000 moderated Z-score values from LINCS L1000 data for measured genes at the highest dose in MCF7 cells. Compounds are presented as grouped by the pathway they target. Red point corresponds to fluvestrant. **C)** Similar to panel A, but for fluvestrant in MCF7 cells and at the highest dose (10 μM). **D)** Similar to panel B, but for A549 cells. Red point corresponds to triamcinalone acetonide. **E)** Similar to panel A, but for triamcinalone acetonide in A549 cells and at the highest dose (10 μM).


A
A549
B
K562
C
MCF7
0
Proliferation Index
4
D
A549
E
K562
F
MCF7
Density
0.0
0.2
0.4
0.6
0.8
0
1
2
3
4
5
Proliferation Index
G
A549
H
K562
I
MCF7
Viability
0.00
0.25
0.50
0.75
1.00
10-20%
20-30%
30-40%
40-50%
50-60%
70-80%
80-90%
90-100%
Percentage of low proliferation index cells per drug-dose combination
J
A549
K562
MCF7
-log10(q_value)
0
5
10
15
20
25
-2
0
2
log 2 fold change in expression (high proliferation vs. low proliferation cells)

Supplementary Figure 11. Single cell measurements reveal variation in proliferation status in vehicle treated cell and across each dose of each drug. A-C) UMAP projection of A549 **(A)**, K562 **(B)** and MCF7 **(C)** colored by proliferation index. High proliferation index indicates an increase in the aggregate expression of transcripts that are markers for G1/S phase or G2/M phase (*43*). **(D-F)** Density plot of cell cycle distribution for compound-treated cells (blue fill) or vehicle-treated cells (red line). Grey line indicates cutoff used to distinguish proliferating cells (greater than cutoff) vs. non-proliferating cells (less than cutoff). **G-I)** Relationship between the percentage of cells designated as low proliferation at each dose of each drug (x-axis) versus the median estimated viability of that combination (y-axis). Each black point corresponds to cells treated with the same dose of a given drug. Red points correspond to vehicle treatment. **J)** Volcano plot depicting the $\log_2$ fold change for significant (q value < 0.01) differentially expressed genes between high and low fractions of vehicle treated cells.

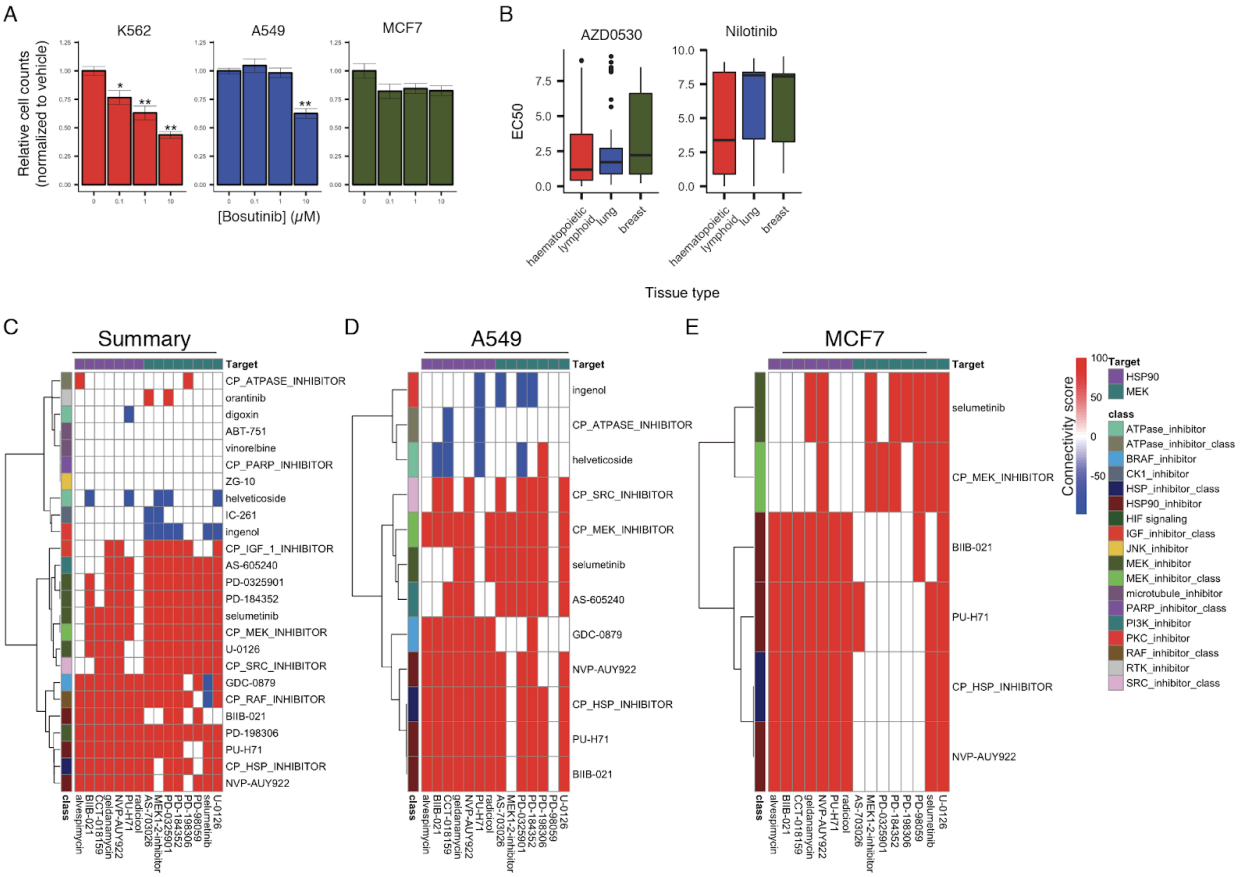


Supplementary Figure 12. Single cell measurements enable estimation of proliferation status and viability across drug-dose combinations. Heatmap depicting estimates of relative proliferation rate, the percentage of cells exhibiting low proliferation index, and the estimated viability for each compound (row) at each dose (column) pair.

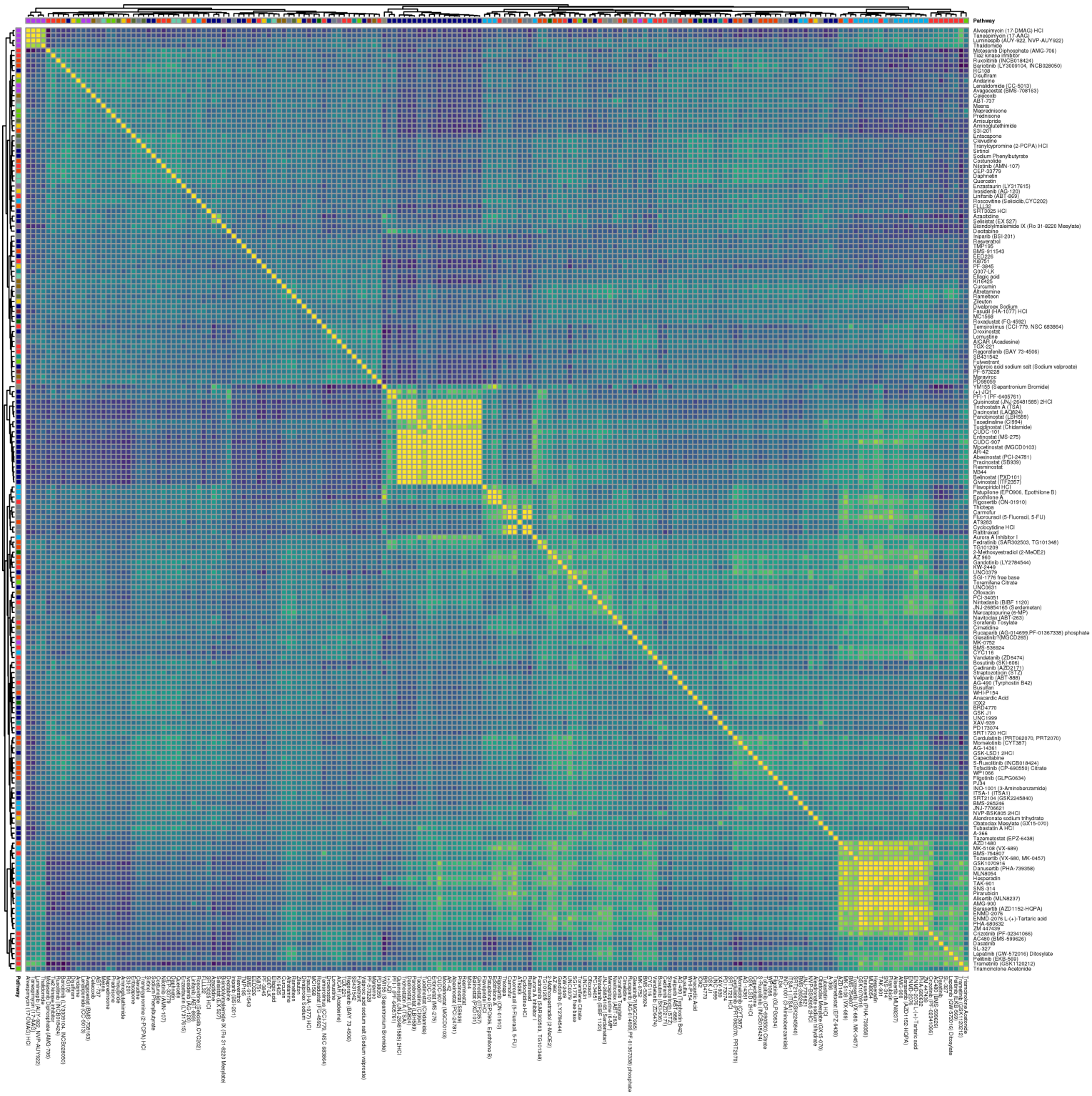


Supplementary Figure 13. sci-Plex enables the dissection of proliferating and non-proliferating cell populations. A) Schematic depicting how changes in cellular state (top) and changes in the relative frequency of subpopulations (bottom) look identical upon subjecting the sample to aggregate measures such as bulk RNA-seq. Adapted from ref (*14*). **B,C)** Pearson correlations between dose-dependent effect sizes estimated from high vs. low proliferation index cells for each cell line (panel B) and drug class (panel C). **D)** Per-gene effect sizes estimated

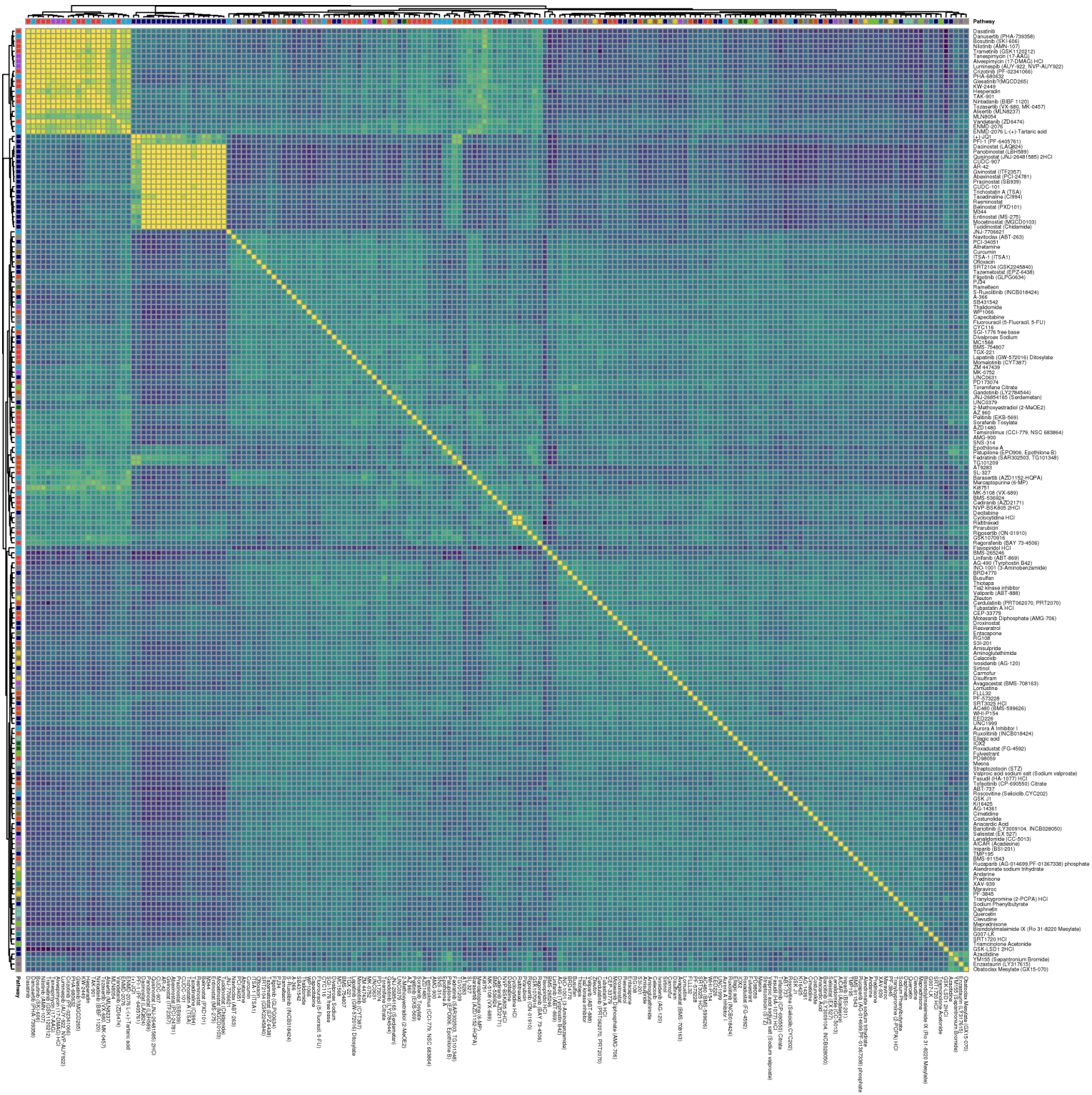
from high (β_{dh}) vs. low (β_{dl}) proliferation index cells for 4 selected compounds. Effect sizes are expressed as $\log_2$ transformed fold changes over intercept. Four classes of genes are shown: those significant in only high proliferation index cells (green); only low proliferation index cells (purple); both high and low cells, and with concordant effect estimates (red); both high and low cells, but with discordant effect estimates (blue). A drug had concordant dose-dependent effects on gene h in high cells (β_{dh}) and low cells (β_{dl}) when $|\beta_{dh} - \beta_{dl}|$ was less than 10 percent of $\frac{1}{2}(|\beta_{dh}| + |\beta_{dl}|)$. Black line indicates y = x.



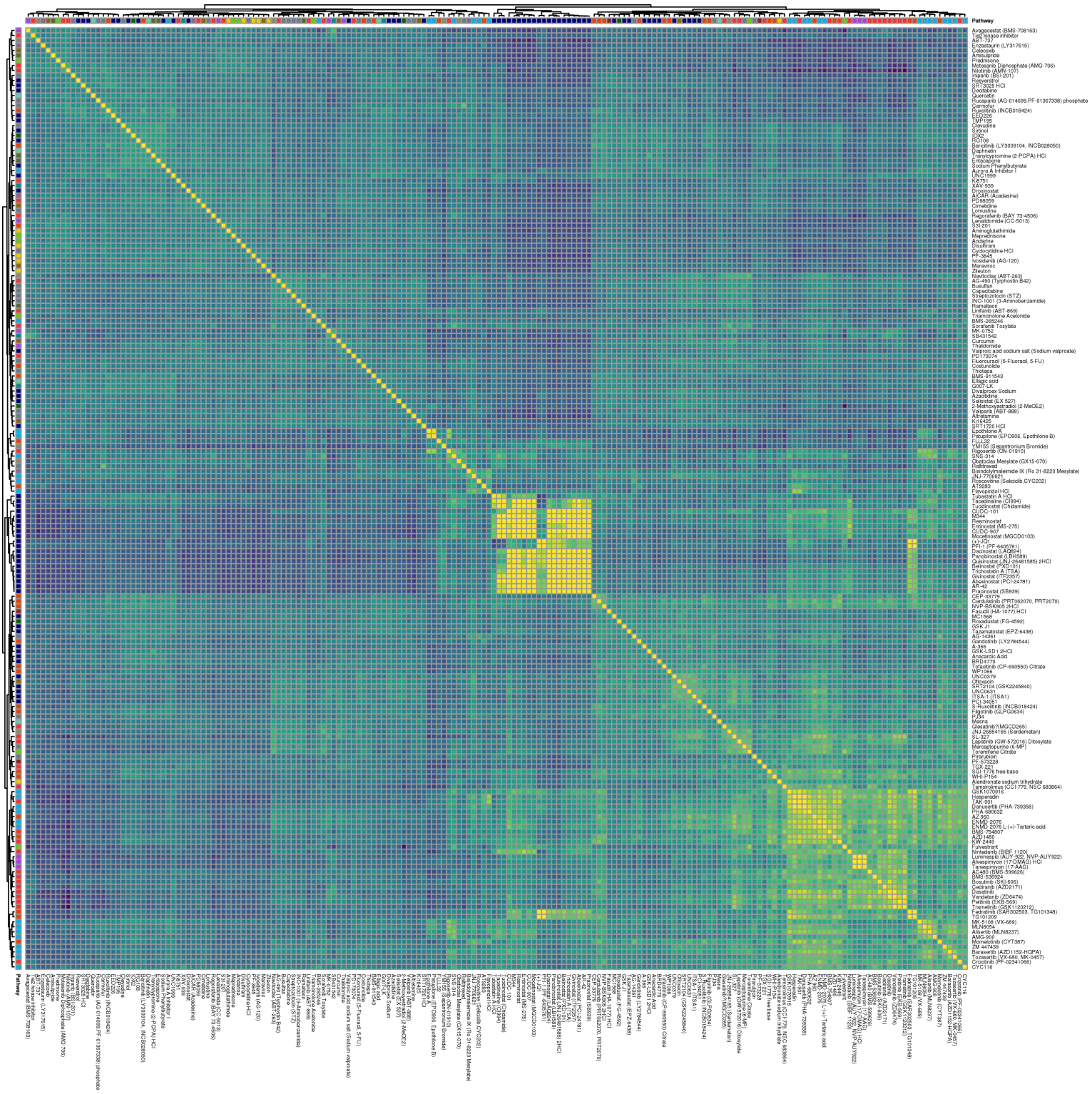
Supplementary Figure 14. sci-Plex screen identifies viability and expression signatures that are reproducible across validation experiments and orthogonal datasets. A) Cell count viability estimates for K562 (red), A549 (blue) and MCF7 (green) cells exposed to vehicle or increasing doses of the Src/Abl inhibitor bosutinib (n = 6 culture replicates, Wilcoxon rank sum test). For each cell line, cell count values were normalized to the mean cell counts value of vehicle control treated cells. Error bars denote standard error of the mean, n = 8. **B)** EC50 values for cell lines of hematopoietic and lymphoid, lung and breast tissue origin, for which viability estimates are available from the Cancer Cell Line Encyclopedia (CCLE), exposed to the Abl inhibitors AZD0530 (left panel) or nilotinib (right panel). **C-E)** Top connectivity scores (a measure that summarizes similarities between transcriptional signatures induced by different drugs (*11, 12*)) for MEK and HSP inhibitors from the CMAP database across all cell lines (summary, panel C) or for A549 (panel D) and MCF7 (panel E) cells individually. A connectivity score cutoff of +/- 90 was applied as in (*11*).



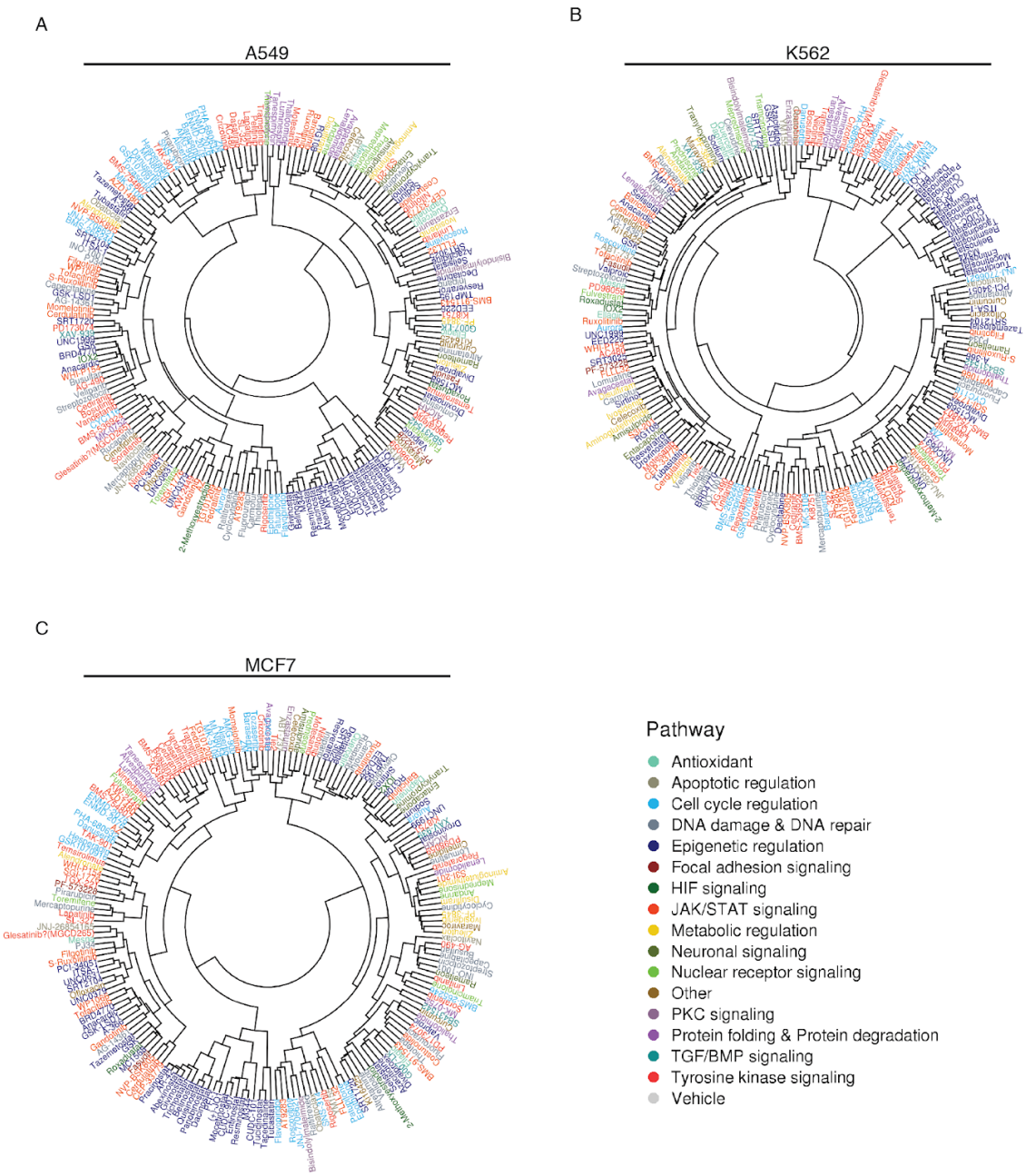
Supplementary Figure 15. Correlation of compound-driven molecular signatures for A549 cells identified in sci-Plex screen. Heatmap depicts the Pearson correlation of beta coefficients across dose-dependent differentially expressed genes for every pairwise combination of compounds screened. To aid in visualization Pearson correlations were capped at 0.6.



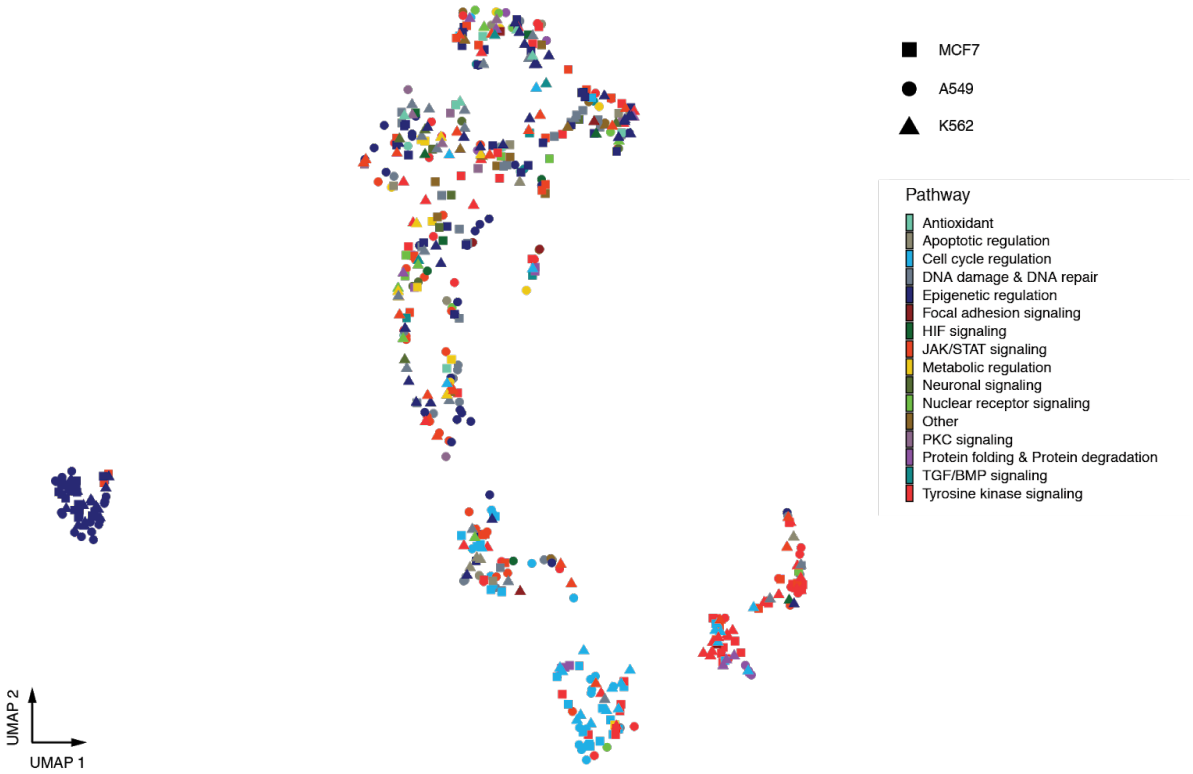
Supplementary Figure 16. Correlation of compound-driven molecular signatures for K562 cells identified in sci-Plex screen. Heatmap depicts the Pearson correlation of beta coefficients across dose-dependent differentially expressed genes for every pairwise combination of compounds screened. To aid in visualization Pearson correlations were capped at 0.6.



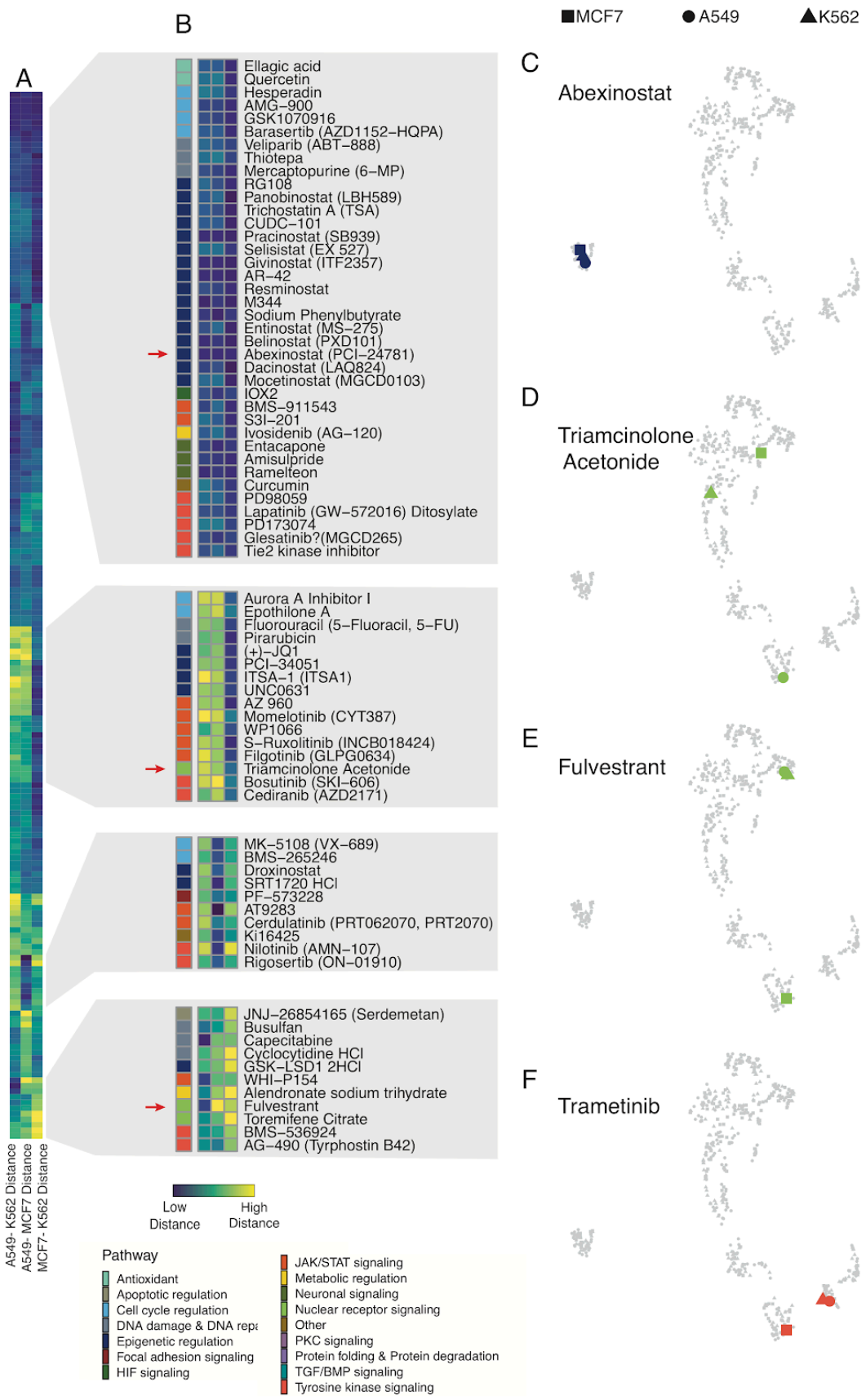
Supplementary Figure 17. Correlation of compound-driven molecular signatures for MCF7 cells identified in sci-Plex screen. Heatmap depicts the Pearson correlation of beta coefficients across dose-dependent differentially expressed genes for every pairwise combination of compounds screened. To aid in visualization Pearson correlations were capped at 0.6.



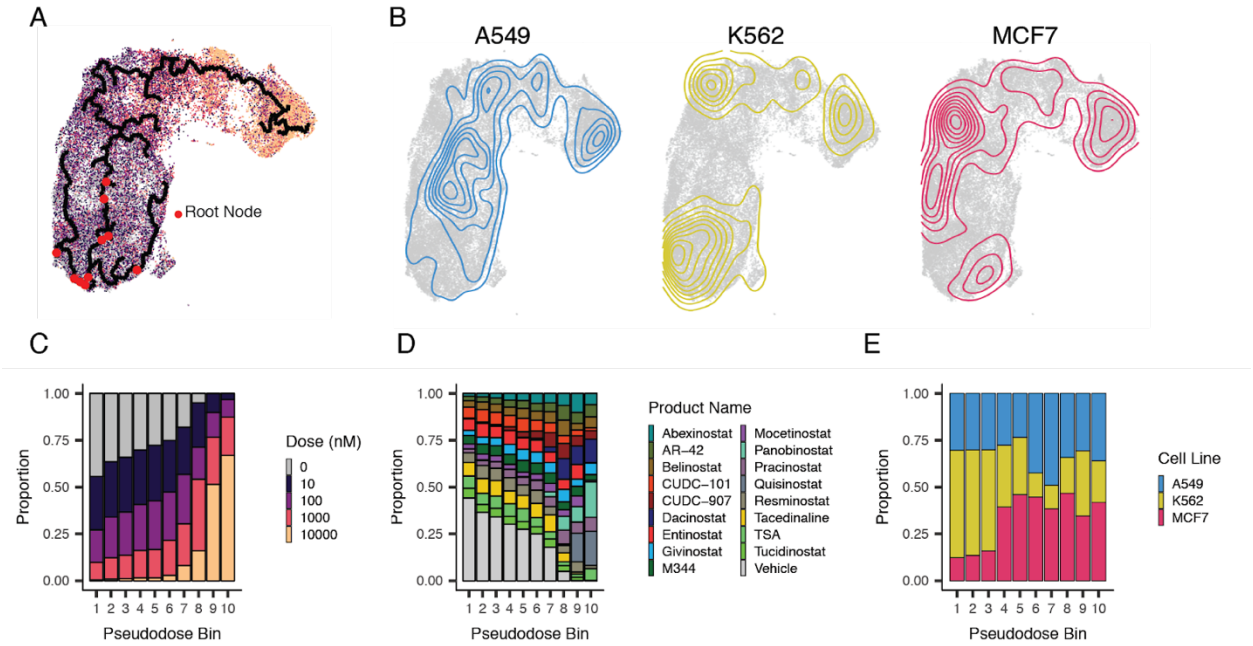
Supplementary Figure 18. **Clustergrams of the correlation of compound-driven molecular signatures.** Clustergrams depicting the Pearson correlation of beta-coefficients across dose-dependent differentially expressed genes for every pairwise combination of compounds screened for A549 (**A**), K562 (**B**) and MCF7 (**C**) cells. Compounds names are colored by the pathway targeted.



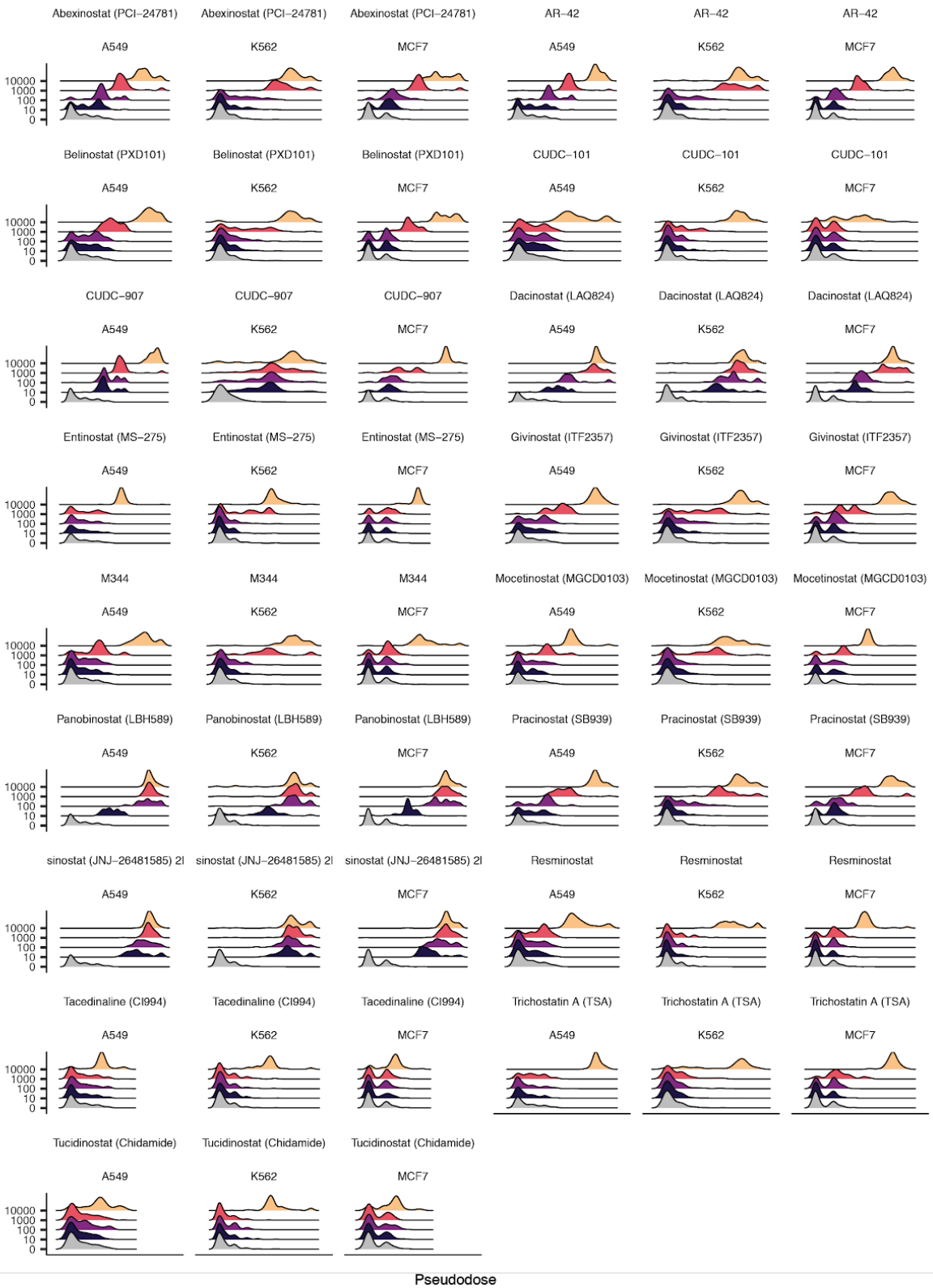
Supplementary Figure 19. UMAP embedding of drugs based on their dose-dependent effects on each gene's expression. Each drug was provided to UMAP as a vector of the effect estimates (β_d, see **Methods**) for all genes. Point shape corresponds to cell type and color corresponds to compound class.


A
B
C
D
E
F
MCF7
A549
K562
Ellagic acid
Quercetin
Hesperadin
AMG−900
GSK1070916
Barasertib (AZD1152−HQPA)
Veliparib (ABT−888)
Thiotepa
Mercaptopurine (6−MP)
RG108
Panobinostat (LBH589)
Trichostatin A (TSA)
CUDC−101
Pracinostat (SB939)
Selisistat (EX 527)
Givinostat (ITF2357)
AR−42
Resminostat
M344
Sodium Phenylbutyrate
Entinostat (MS−275)
Belinostat (PXD101)
Abexinostat (PCI−24781)
Dacinostat (LAQ824)
Mocetinostat (MGCD0103)
IOX2
BMS−911543
S3I−201
Ivosidenib (AG−120)
Entacapone
Amisulpride
Ramelteon
Curcumin
PD98059
Lapatinib (GW−572016) Ditosylate
PD173074
Glesatinib?(MGCD265)
Tie2 kinase inhibitor
Aurora A Inhibitor I
Epothilone A
Fluorouracil (5−Fluoracil, 5−FU)
Pirarubicin
(+)−JQ1
PCI−34051
ITSA−1 (ITSA1)
UNC0631
AZ 960
Momelotinib (CYT387)
WP1066
S−Ruxolitinib (INCB018424)
Filgotinib (GLPG0634)
Triamcinolone Acetonide
Bosutinib (SKI−606)
Cediranib (AZD2171)
MK−5108 (VX−689)
BMS−265246
Droxinostat
SRT1720 HCl
PF−573228
AT9283
Cerdulatinib (PRT062070, PRT2070)
Ki16425
Nilotinib (AMN−107)
Rigosertib (ON−01910)
JNJ−26854165 (Serdemetan)
Busulfan
Capecitabine
Cyclocytidine HCl
GSK−LSD1 2HCl
WHI−P154
Alendronate sodium trihydrate
Fulvestrant
Toremifene Citrate
BMS−536924
AG−490 (Tyrphostin B42)
Abexinostat
Triamcinolone Acetonide
Fulvestrant
Trametinib
A549- K562 Distance
A549- MCF7 Distance
MCF7- K562 Distance
Low Distance
High Distance
Pathway
Antioxidant
Apoptotic regulation
Cell cycle regulation
DNA damage & DNA repa
Epigenetic regulation
Focal adhesion signaling
HIF signaling
JAK/STAT signaling
Metabolic regulation
Neuronal signaling
Nuclear receptor signaling
Other
PKC signaling
Protein folding & Protein degradation
TGF/BMP signaling
Tyrosine kinase signaling

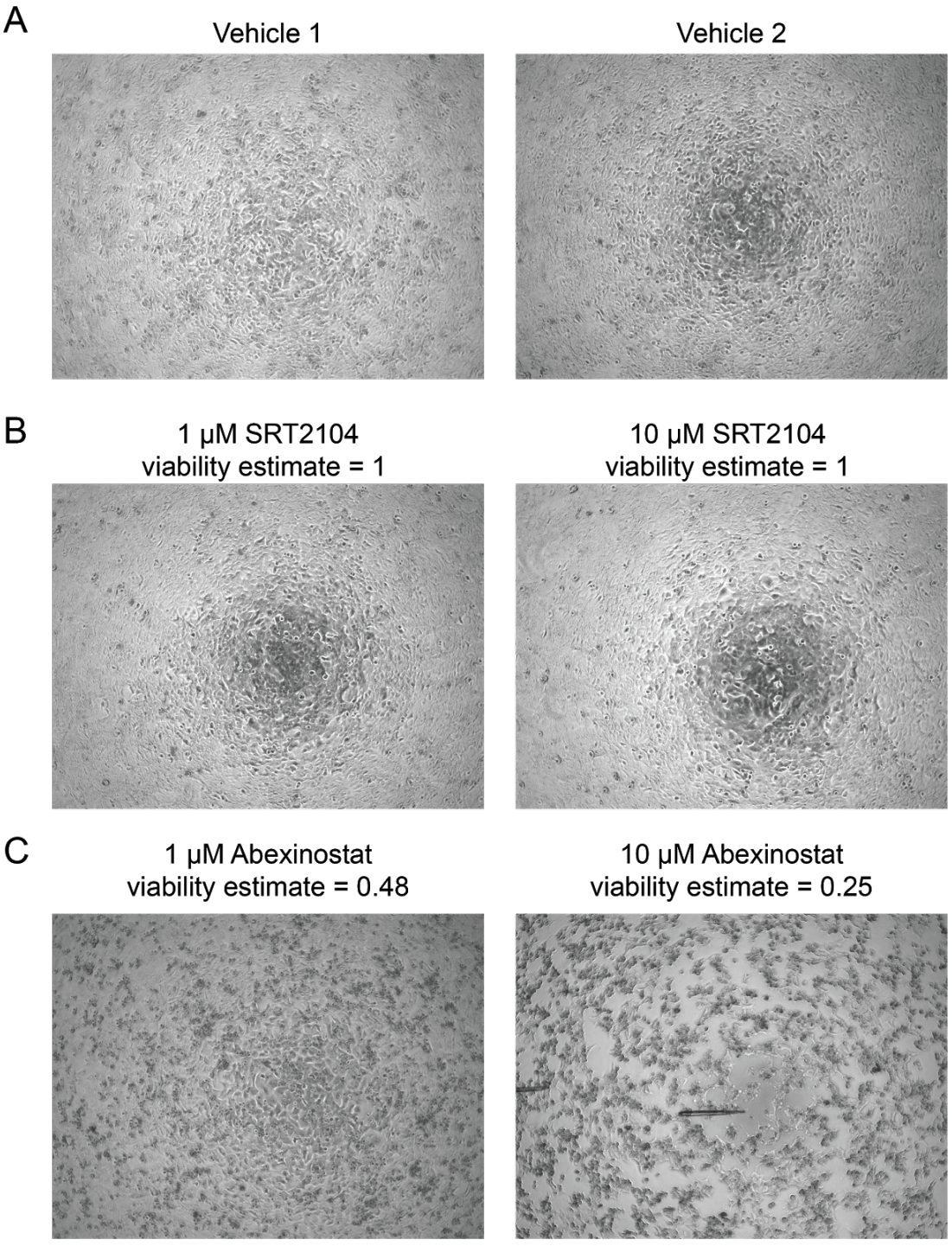
Supplementary Figure 20. Pairwise distances between PCA embeddings of drugs based on their dose-dependent effects. A) Heatmap of pairwise distances between two cell types (columns) for a given drug (rows) in PCA reduced dimensional space. Hierarchically clustered to visualize cell type-specific responses to each drug. **B)** Insets of highlighted portions of the heatmap with pathway annotation shown to the left. Specific compounds highlighted with a red arrow are shown to the right **(C-E)** as UMAP embeddings. **F)** Trametinib treated cell lines are highlighted to illustrate colocalization of A549 and K562. Colored points correspond to labeled compound and all other drugs are shown in gray. Shape encodes the cell line from which each effect profile was captured (squares: MCF7; triangles: K562; circle: A549).



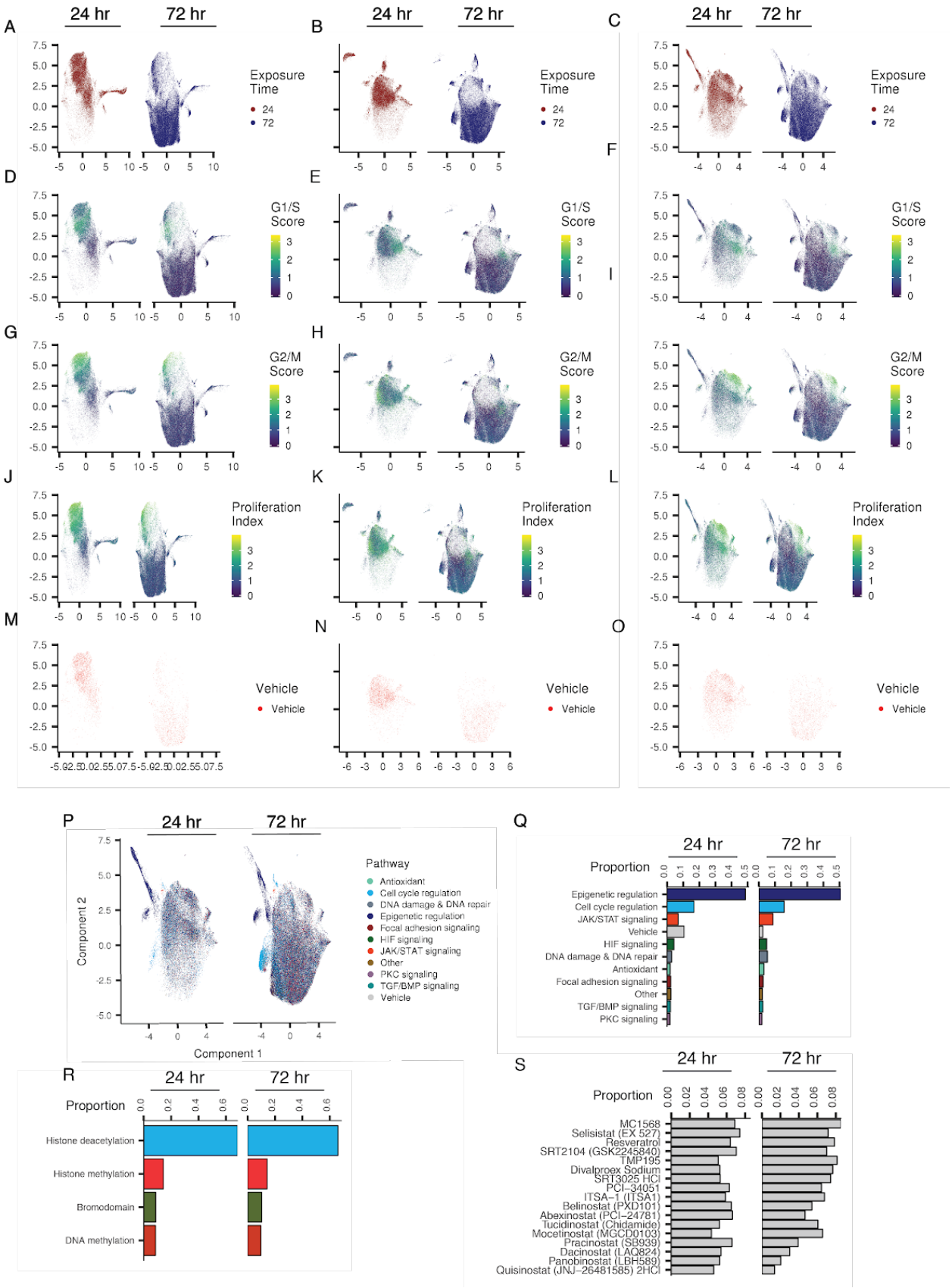
Supplementary Figure 21. HDAC inhibitor-treated cell types align and enable joint pseudodose trajectory reconstruction. A) UMAP embedding highlighting the reconstructed pseudodose trajectory over the mutual nearest neighbor-aligned HDAC inhibitor and vehicle treated cells. Root nodes (red points) were chosen as nodes in the principal graph that had over 50% of their nearest neighbors annotated as vehicle treated cells. **B)** Distribution of each cell line within the embedding. **C)** Barplot displaying the fraction of each pseudodose bin occupied by cells treated at each dose. **D)** Barplot displaying the fraction of each pseudodose bin occupied by cells treated with each compound. **E)** Proportion within each pseudodose bin corresponding to each cell line.



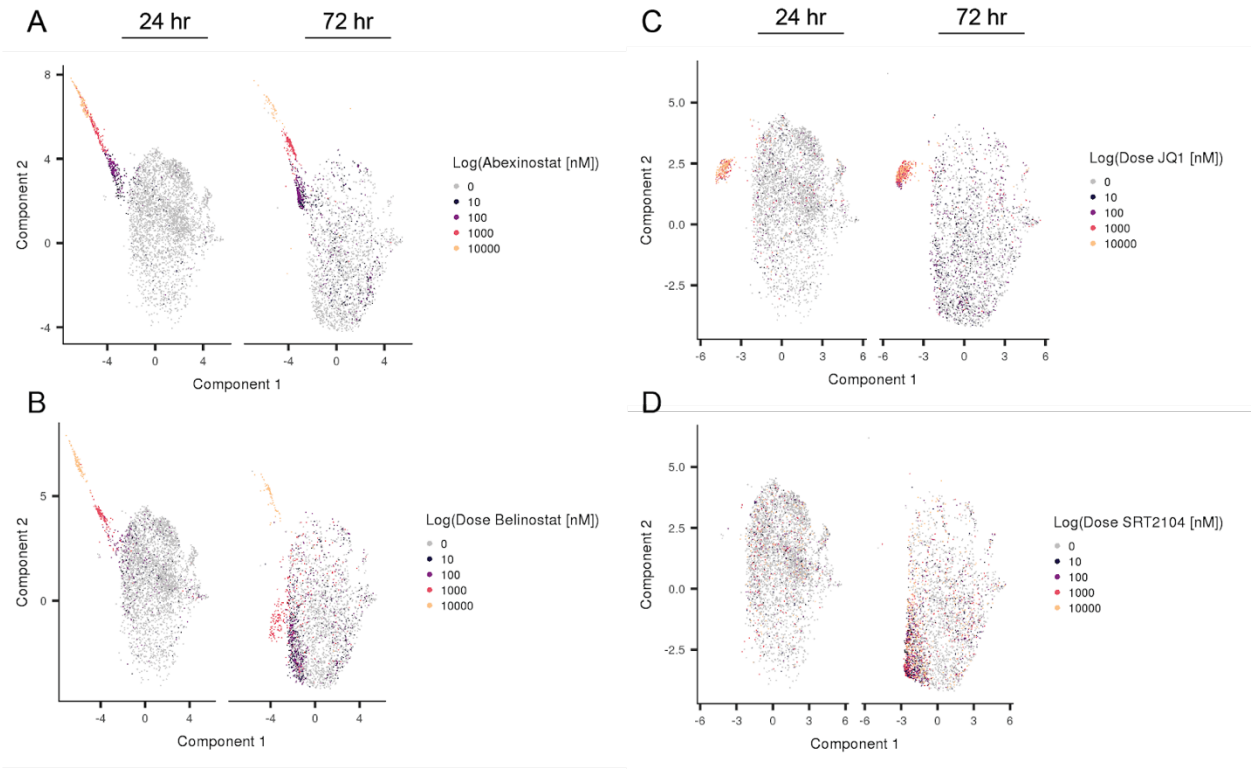
Supplementary Figure 22. Ridge plots display the distribution of cells along pseudodose for each HDAC inhibitor and dose combination for compounds that localized to the HDAC trajectory.



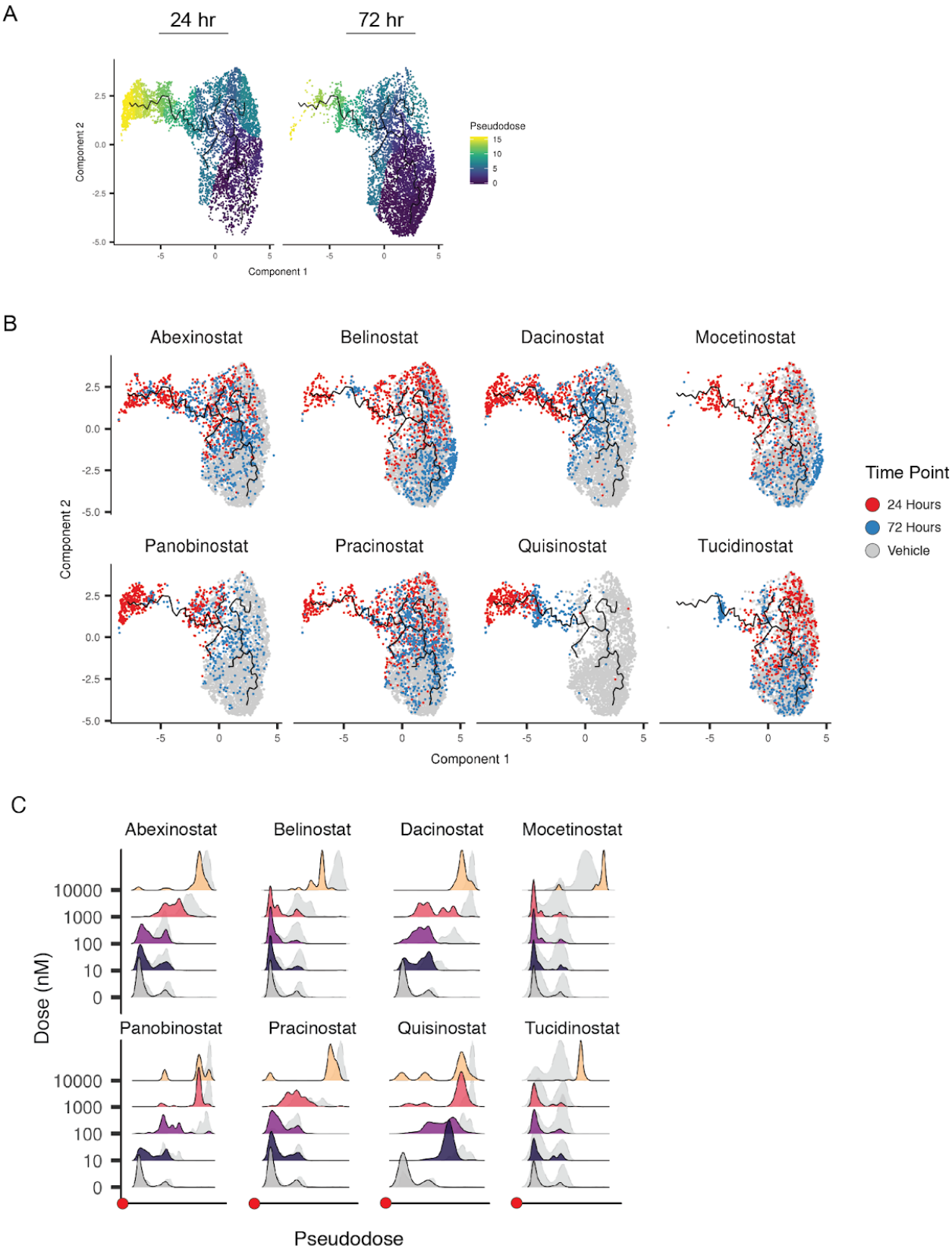
Supplementary Figure 23. Contact inhibition of cell proliferation 72 hours post drug exposure. Representative brightfield images of A549 cells exposed to vehicle (**A**) or the specified dose of the SIRT1 activator SRT2104 (**B**) or the HDAC inhibitor Abexinostat (**C**). Viability estimates as determined by recovered cell counts for each drug/dose combination normalized to cell counts of vehicle control wells.


A
B
C
D
E
F
G
H
I
J
K
L
M
N
O
24 hr
72 hr
Exposure Time
24
72
G1/S Score
G2/M Score
Proliferation Index
Vehicle
P
Component 1
Component 2
Pathway
Antioxidant
Cell cycle regulation
DNA damage & DNA repair
Epigenetic regulation
Focal adhesion signaling
HIF signaling
JAK/STAT signaling
Other
PKC signaling
TGF/BMP signaling
Q
Proportion
R
Histone deacetylation
Histone methylation
Bromodomain
DNA methylation
S
MC1568
Selisistat (EX 527)
Resveratrol
SRT2104 (GSK2245840)
TMP195
Divalproex Sodium
SRT3025 HCl
PCI-34051
ITSA-1 (ITSA1)
Belinostat (PXD101)
Abexinostat (PCI-24781)
Tucidinostat (Chidamide)
Mocetinostat (MGCD0103)
Pracinostat (SB939)
Dacinostat (LAQ824)
Panobinostat (LBH589)
Quisinostat (JNJ-26481585) 2HCl

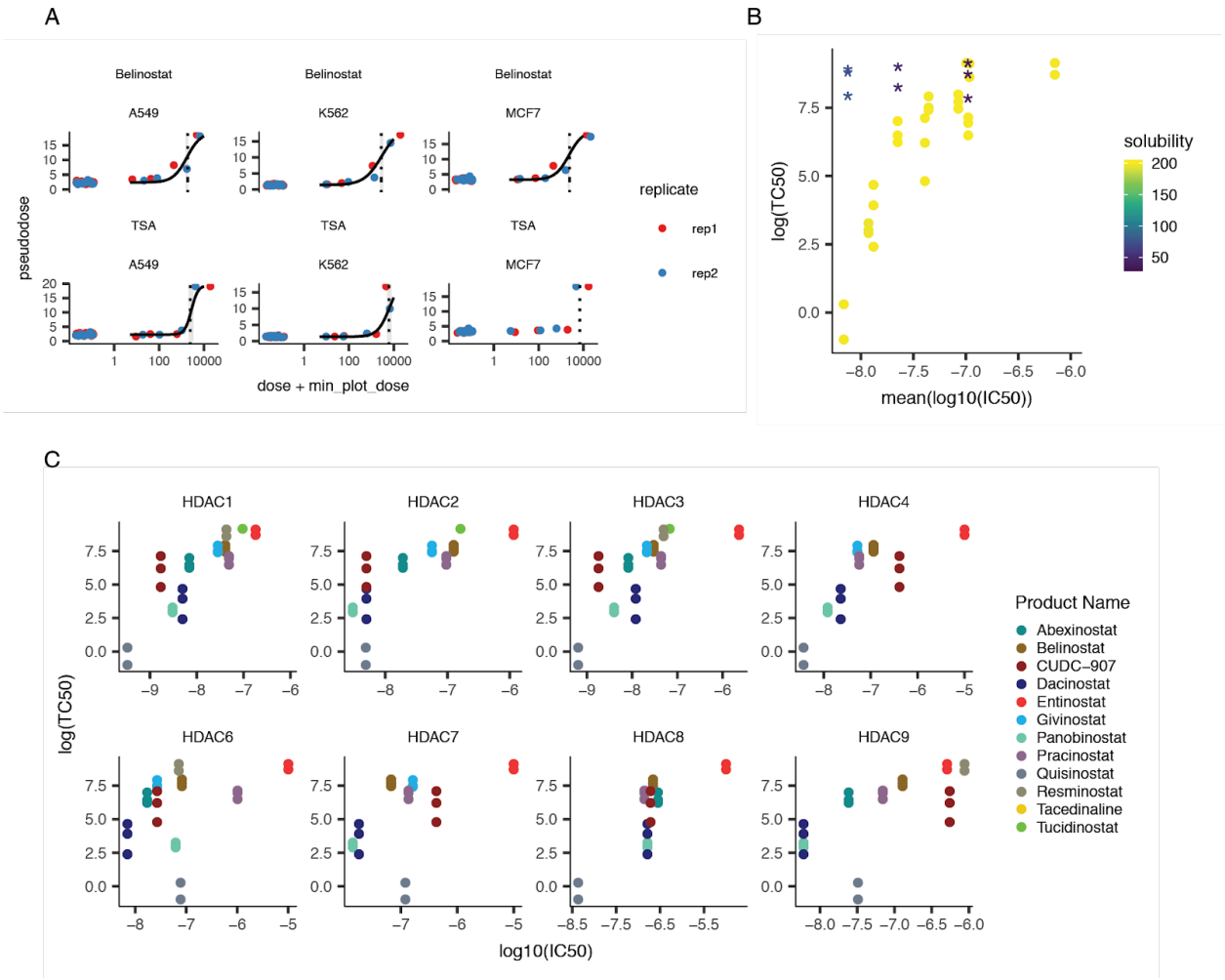
Supplementary Figure 24. Aligning A549 cells at 24 and 72 hours after treatment reveals time-dependent responses to diverse small molecules. (**A-C)** UMAP embedding of A549 cells at 24 and 72 hours post treatment in the absence of a correction for differences in viability and proliferation **(A)**, after linear transformation of the data to account for changes in proliferation index and viability **(B)** and after mutual nearest neighbor based alignment of data after linear transformation **(C)**. Cells are colored by the time point at which they were collected. **(D-F)** UMAP embeddings as in panels A-C with cells colored by the aggregated normalized expression score of G1/S marker genes. **(G-I)** UMAP embeddings as in panels A-C with cells colored by the aggregated normalized expression score of G2/M marker genes. **(J-L)** UMAP embeddings as in panels A-C with cells colored by proliferation index. **(M-O)** UMAP embeddings as in panels A-C only visualizing cells treated with vehicle control. **(P)** UMAP embeddings from panel C with cells colored as to the pathway targeted by the treatment to which they were exposed. **(Q)** Proportion of cells broken up by pathway targeted. Note that only a subset of our 188 compounds across a limited number of pathways were tested at 72 hours. **(R)** Proportion of cells broken up by the activity targeted by treatment with epigenetic regulation compounds. **(S)** Proportion of cells broken up by HDAC compound.



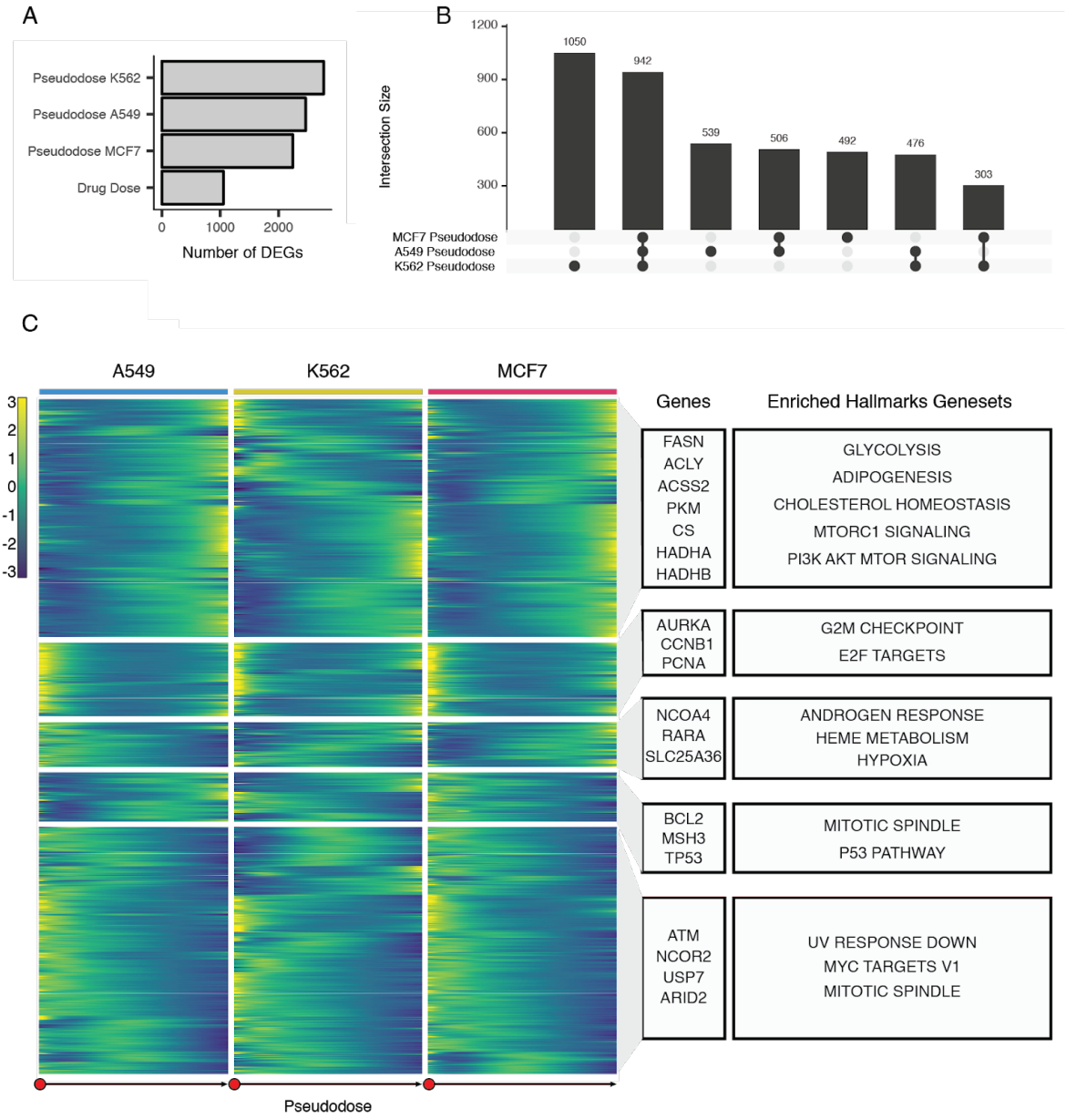
Supplementary Figure 25. Bromodomain inhibition, sirtuin activation, and histone deacetylase inhibition induce characteristic transcriptomic responses. (A-D) UMAP embedding of MNN aligned A549 cells 24 and 72 hours after treatment with the pan-HDAC inhibitors abexinostat (**A**) or belinostat (**B**), the bromodomain inhibitor JQ1 (**C**), and the SIRT1 activator SRT2104 (**D**). Cells are colored by the dose to which each cell was exposed.


A
24 hr
72 hr
Component 2
Component 1
Pseudodose
15
10
5
0
B
Abexinostat
Belinostat
Dacinostat
Mocetinostat
Panobinostat
Pracinostat
Quisinostat
Tucidinostat
Component 2
Component 1
Time Point
24 Hours
72 Hours
Vehicle
C
Abexinostat
Belinostat
Dacinostat
Mocetinostat
Panobinostat
Pracinostat
Quisinostat
Tucidinostat
Dose (nM)
10000
1000
100
10
0
Pseudodose

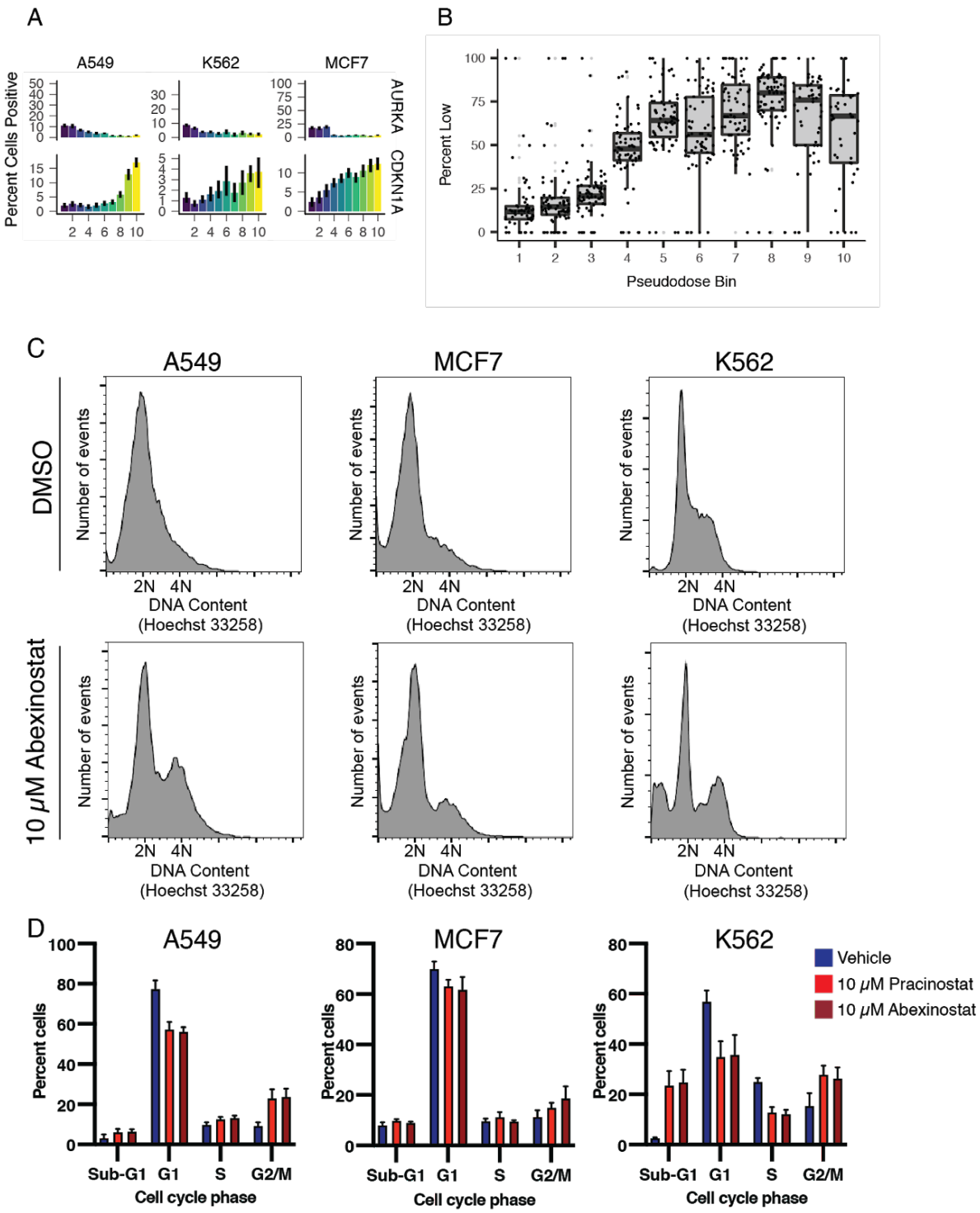
Supplementary Figure 26. The heterogeneous response to the majority of HDAC inhibitors does not appear to be driven by cellular asynchrony. A) Aligned UMAP embeddings of cells exposed to vehicle HDAC inhibitors for 24 or 72 hours. Cells are colored by their progression along pseudodose. **B)** Aligned UMAP embeddings of cells exposed to vehicle (grey cells) or the labeled HDAC inhibitor for 24 (red cells) or 72 (blue cells) hours. **C)** Ridge plots displaying the density of HDAC inhibitor-exposed A549 cells along an aligned pseudodose trajectory. Results are displayed for the 8 HDAC inhibitors that were assayed at both 24 and 72 hours. Gray and color filled lines denote cells exposed with inhibitors for 24 or 72 hours, respectively.



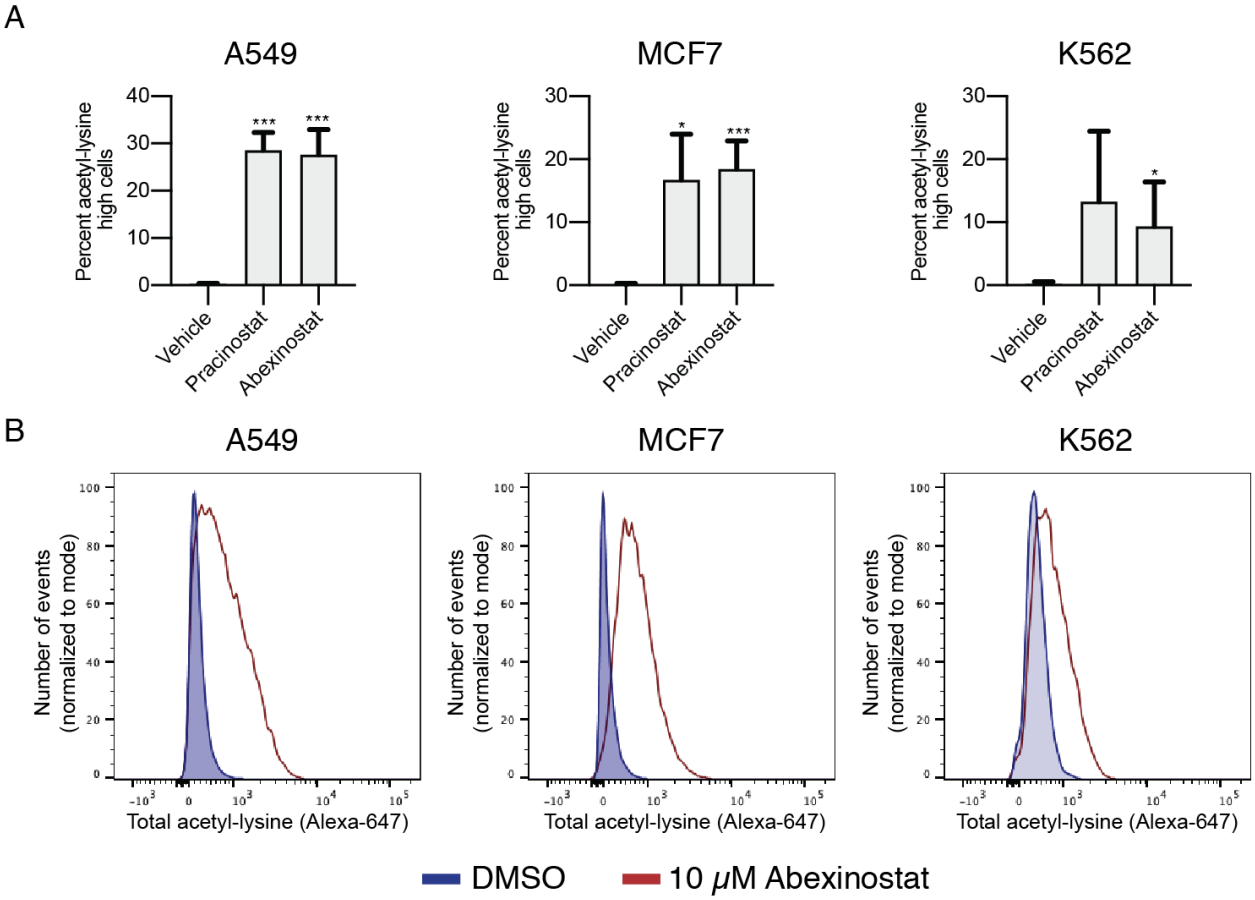
Supplementary Figure 27. Transcriptional trajectory of HDAC inhibitor-treated cells corresponds to *in vitro* IC50 measurements. A) Pseudodose response curves were fit for each compound and each cell line using the drc R package. The mean position of each dose along the pseudodose trajectory was used as the response. Two illustrative examples for belinostat (top) and trichostatin A (TSA) (below) are shown. Dotted vertical lines illustrate the transcriptional EC50 (TC50) for each compound in each cell line. Shaded gray area denotes the 95% confidence intervals for each TC50 estimate. B) Plot displaying aggregate *in vitro* measured mean of $\log_{10}$(IC50 [M]) versus log(TC50) colored by solubility supplied by Selleckchem Chemicals. Points displayed as (*) were not used for fits. C) $\log_{10}$(IC50 [M]) versus log(TC50) for each HDAC isoform. Each point is colored by the HDAC inhibitor used.



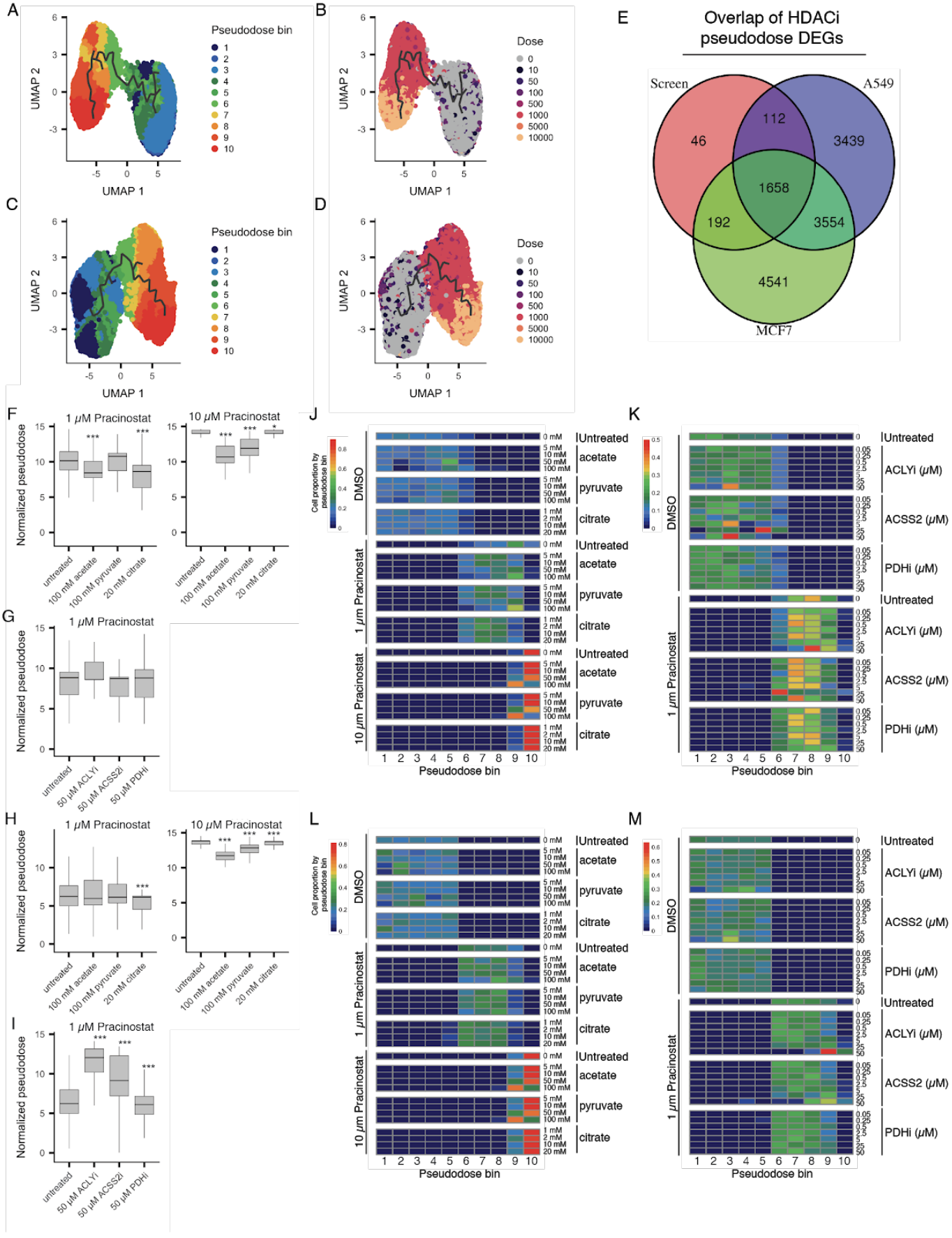
Supplementary Figure 28. Linear models identify pseudodose-dependent modules of proliferation and metabolism. A) Barplot of the total number of significant dose-dependent and pseudodose-dependent DEGs (FDR < 0.05). **B)** Upset plot displaying the intersections of significant pseudodose-dependent DEGs between the three cell types. **C)** Pseudodose heatmap depicting 4,308 genes that varied significantly as a function of pseudodose. Each row corresponds to the expected expression for a gene in the three cell lines as fit by the model described in the 'Differential expression analysis' section of the Methods. Genes (rows) were scaled and standardized within each cell line before joining the three matrices and performing hierarchical clustering. Clusters from hierarchical clustering were then used as an input into GSAhyper using the Hallmarks geneset collection. Select genes and genesets characterizing each cluster are shown (right).



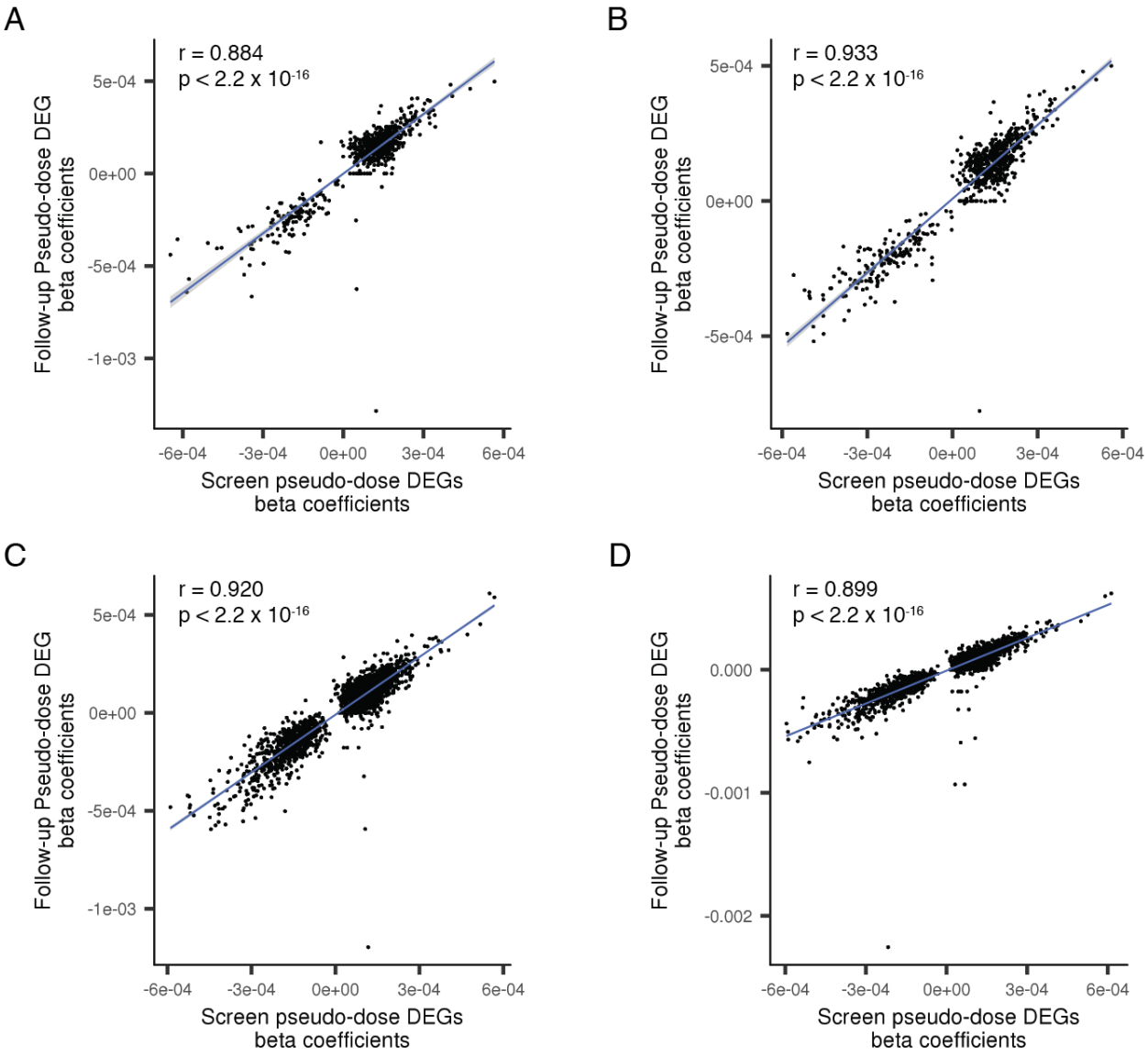
Supplementary Figure 29. HDAC inhibitor treatment induces cell cycle arrest in all three cell lines. A) Percentage of cells expressing RNA for AURKA and CDKN1A across pseudodose bins. Black bars denote the bootstrapped 95% confidence interval. **B)** Boxplots depicting the percentage of cells in the low proliferation fraction in at a given drug dose across pseudodose bins. **C)** DNA content analysis of the three cell lines upon treatment with DMSO (top) or 10μM abexinostat (bottom). **D)** Quantification of flow cytometry data depicting the number of cells in each DNA content category.



Supplementary Figure 30. HDAC inhibitor exposure leads to sequestration of acetate in the form of acetylated lysines. A) Quantification of flow cytometry measurements of total cellular acetylated lysines in A549 (left panel), MCF7 (middle panel) and K562 (right panel) cells exposed to 10 µM pracinostat, 10 µM p abexinostat or vehicle control. Error bars denote standard deviation of the mean (Wilcoxon rank sum test, n = 3 culture replicates, * $p < 0.05$, *** $p < 0.005$). **B**) Representative flow cytometry histograms for the experiment quantified in panel A. Blue shaded regions and red lines correspond to DMSO vehicle control and 10 µM abexinostat, respectively.


A
B
C
D
E
F
G
H
I
J
K
L
M
UMAP 1
UMAP 2
Pseudodose bin
Dose
Overlap of HDACi pseudodose DEGs
Screen
A549
MCF7
46
112
3439
1658
192
3554
4541
1 μM Pracinostat
10 μM Pracinostat
Normalized pseudodose
untreated
100 mM acetate
100 mM pyruvate
20 mM citrate
50 μM ACLYi
50 μM ACSS2i
50 μM PDHi
Cell proportion by pseudodose bin
DMSO
1 μm Pracinostat
10 μm Pracinostat
Untreated
acetate
pyruvate
citrate
ACLYi (μM)
ACSS2 (μM)
PDHi (μM)
Pseudodose bin

Supplementary Figure 31. Supplementation with acetyl-CoA precursors decrease, while inhibition of enzymes that replenish acetyl-coA pools exacerbate, progression along the HDAC inhibitor pseudodose trajectory. A-D) UMAP embeddings of A549 (panels A and B) and MCF7 (panels C and D) single cell transcriptomes after exposure to the HDAC inhibitors pracinostat or abexinostat, in the presence or absence of acetyl-CoA precursors or inhibitors to enzymes that replenish acetyl-CoA pools. UMAP were constructed from cells from all conditions in the experiment. Cells are colored by pseudodose bin (panels A and C) or dose (panels B and D). **E)** Venn diagram of the overlap of differentially expressed genes across trajectories between or original HDACi trajectory vs. A549 or MCF7 HDACi trajectories from this new experiment. **F,H)** Boxplots of pseudodose estimates for select conditions of cells exposed to 1 or 10 μM pracinostat with or without co-treatment with acetyl-coA precursors for A549 (panel H) or MCF7 (panel L) cells. Values are normalized to vehicle treated cells. Wilcoxon rank sum test. **G,I)** Boxplots of pseudodose estimates for select conditions of cells exposed to vehicle and pracinostat with or without co-treatment with acetyl-coA precursors for A549 (panel I) or MCF7 (panel M) cells. Values were normalized to vehicle treated cells. Wilcoxon rank sum test. **J,L)** Heatmaps depicting the fraction of cells per pseudodose bin for cells exposed to various acetyl-coA precursors in pracinostat-exposed A549 (**F**) or MCF7 (**J**) cell. **K,M)** Heatmaps depicting the fraction of cells per pseudodose bin for cells exposed to various inhibitors targeting enzymes that replenish acetyl-coA pools in pracinostat-exposed A549 (panel G) and MCF7 (panel K) cells.



Supplementary Figure 32. Correlation of effect sizes between differentially expressed genes post-HDAC inhibition from original screen vs. new experiment. A-B) Correlation of effect size estimates (beta coefficients) for differentially expressed genes between vehicle control and 10 μM abexinostat (panel A) or 10 μM pracinostat (panel B) for A549 cells. **C-D**) Correlation of effect size estimates (beta coefficients) for differentially expressed genes between vehicle control and 10 μM abexinostat (panel C) or 10 μM pracinostat (panel D) for MCF7 cells. X-axes correspond to large-scale sci-Plex experiment. Y-axes correspond to targeted follow-up sci-Plex experiment.

Experiment	sci-RNA-seq protocol used	Cells Profiled	RNA Treshold (# UMIs)	Hash Oligo Treshold (# UMIs)	Hash Oligo Enrichment Ratio (UMIs Top Rank Oligo / UMIs Second Rank Oligo)
Barnyard	2-level	3024	1000	10	15
Proof-of-concept Screen	2-level	12,435	1000	30	10
Large Screen	3-level	649,220	500	5/5 (Well Oligo and Plate	5/5 (Well Oligo and Plate Oligo)
HDACi phenocopy/rescue	3-level	72,966	500	5	5

Supplementary Table 1. Summary of single cell sequencing experiments performed and thresholds used in analysis of the data. Threshold values chosen manually based on distribution of molecules observed. Cells that failed to meet thresholds were not included in analyses.

Supplementary Tables Available Online:

Supplementary Table 2. Differential expression analysis of A549 cells treated with Nutlin-3a, SAHA, Dexamethasone, and BMS345541. Differential gene expression test results, model components, and hypothesis testing statistics for each gene tested for each compound.

Supplementary Table 3. Metadata of molecular compounds profiled. Compound specific metadata include molecular weight, SMILES string and CAS number. Experiment specific metadata include pathway annotations, dose, and identifying hash combination.

Supplementary Table 4. DEGs detected per compound per cell line. The number of significant dose-dependent differentially expressed genes (FDR < 0.05) are listed by cell line and compound profiled.

Supplementary Table 5. Differential expression analysis of A549, K562 and MCF7 cells in response to treatment with one of 188 compounds for 24 hours or 72 hours. Differential gene expression test results, model components, and hypothesis testing statistics for each gene for every compound.

Supplementary Table 6. Dose-response model parameters for transcriptionally derived HDAC inhibitor potencies. Mean pseudodose values were fit as a function of compound concentration using the drc package. Column names identify parameters.

Supplementary Table 7. *In vitro* IC50 measurements for Histone deacetylase (HDAC) inhibitors in the trajectory analysis. IC50 values collected from published manuscripts (PMID listed in table). NAs represent untested isoforms or missing data.

Supplementary Table 8. DEGs detected over consensus HDAC inhibition trajectory. Differential gene expression test results, model components, and hypothesis testing statistics for each gene for every compound.

References and Notes

1. J. R. Broach, J. Thorner, High-throughput screening for drug discovery. *Nature* **384**, 14–16 (1996). Medline

2. D. A. Pereira, J. A. Williams, Origin and evolution of high throughput screening. *Br. J. Pharmacol.* **152**, 53–61 (2007). doi:10.1038/sj.bjp.0707373 Medline

3. D. Shum, C. Radu, E. Kim, M. Cajuste, Y. Shao, V. E. Seshan, H. Djaballah, A high density assay format for the detection of novel cytotoxic agents in large chemical libraries. *J. Enzyme Inhib. Med. Chem.* **23**, 931–945 (2008). doi:10.1080/14756360701810082 Medline

4. C. Yu, A. M. Mannan, G. M. Yvone, K. N. Ross, Y.-L. Zhang, M. A. Marton, B. R. Taylor, A. Crenshaw, J. Z. Gould, P. Tamayo, B. A. Weir, A. Tsherniak, B. Wong, L. A. Garraway, A. F. Shamji, M. A. Palmer, M. A. Foley, W. Winckler, S. L. Schreiber, A. L. Kung, T. R. Golub, High-throughput identification of genotype-specific cancer vulnerabilities in mixtures of barcoded tumor cell lines. *Nat. Biotechnol.* **34**, 419–423 (2016). doi:10.1038/nbt.3460 Medline

5. Z. E. Perlman, M. D. Slack, Y. Feng, T. J. Mitchison, L. F. Wu, S. J. Altschuler, Multidimensional drug profiling by automated microscopy. *Science* **306**, 1194–1198 (2004). doi:10.1126/science.1100709 Medline

6. Y. Futamura, M. Kawatani, S. Kazami, K. Tanaka, M. Muroi, T. Shimizu, K. Tomita, N. Watanabe, H. Osada, Morphobase, an encyclopedic cell morphology database, and its use for drug target identification. *Chem. Biol.* **19**, 1620–1630 (2012). doi:10.1016/j.chembiol.2012.10.014 Medline

7. J. Kang, C.-H. Hsu, Q. Wu, S. Liu, A. D. Coster, B. A. Posner, S. J. Altschuler, L. F. Wu, Improving drug discovery with high-content phenotypic screens by systematic selection of reporter cell lines. *Nat. Biotechnol.* **34**, 70–77 (2016). doi:10.1038/nbt.3419 Medline

8. K. L. Huss, P. E. Blonigen, R. M. Campbell, Development of a Transcreener kinase assay for protein kinase A and demonstration of concordance of data with a filter-binding assay format. *J. Biomol. Screen.* **12**, 578–584 (2007). doi:10.1177/1087057107300221 Medline

9. C. Ye, D. J. Ho, M. Neri, C. Yang, T. Kulkarni, R. Randhawa, M. Henault, N. Mostacci, P. Farmer, S. Renner, R. Ihry, L. Mansur, C. G. Keller, G. McAllister, M. Hild, J. Jenkins, A. Kaykas, DRUG-seq for miniaturized high-throughput transcriptome profiling in drug discovery. *Nat. Commun.* **9**, 4307 (2018). doi:10.1038/s41467-018-06500-x Medline

10. E. C. Bush, F. Ray, M. J. Alvarez, R. Realubit, H. Li, C. Karan, A. Califano, P. A. Sims, PLATE-Seq for genome-wide regulatory network analysis of high-throughput screens. *Nat. Commun.* **8**, 105 (2017). doi:10.1038/s41467-017-00136-z Medline

11. A. Subramanian, R. Narayan, S. M. Corsello, D. D. Peck, T. E. Natoli, X. Lu, J. Gould, J. F. Davis, A. A. Tubelli, J. K. Asiedu, D. L. Lahr, J. E. Hirschman, Z. Liu, M. Donahue, B. Julian, M. Khan, D. Wadden, I. C. Smith, D. Lam, A. Liberzon, C. Toder, M. Bagul, M. Orzechowski, O. M. Enache, F. Piccioni, S. A. Johnson, N. J. Lyons, A. H. Berger, A. F. Shamji, A. N. Brooks, A. Vrcic, C. Flynn, J. Rosains, D. Y. Takeda, R. Hu, D. Davison, J. Lamb, K. Ardlie, L. Hogstrom, P. Greenside, N. S. Gray, P. A. Clemons, S. Silver, X.

Wu, W.-N. Zhao, W. Read-Button, X. Wu, S. J. Haggarty, L. V. Ronco, J. S. Boehm, S. L. Schreiber, J. G. Doench, J. A. Bittker, D. E. Root, B. Wong, T. R. Golub, A next generation connectivity map: L1000 platform and the first 1,000,000 profiles. *Cell* **171**, 1437–1452.e17 (2017). doi:10.1016/j.cell.2017.10.049 Medline

12. J. Lamb, E. D. Crawford, D. Peck, J. W. Modell, I. C. Blat, M. J. Wrobel, J. Lerner, J.-P. Brunet, A. Subramanian, K. N. Ross, M. Reich, H. Hieronymus, G. Wei, S. A. Armstrong, S. J. Haggarty, P. A. Clemons, R. Wei, S. A. Carr, E. S. Lander, T. R. Golub, The Connectivity Map: Using gene-expression signatures to connect small molecules, genes, and disease. *Science* **313**, 1929–1935 (2006). doi:10.1126/science.1132939 Medline

13. M. B. Elowitz, A. J. Levine, E. D. Siggia, P. S. Swain, Stochastic gene expression in a single cell. *Science* **297**, 1183–1186 (2002). doi:10.1126/science.1070919 Medline

14. C. Trapnell, Defining cell types and states with single-cell genomics. *Genome Res.* **25**, 1491–1498 (2015). doi:10.1101/gr.190595.115 Medline

15. S. M. Shaffer, M. C. Dunagin, S. R. Torborg, E. A. Torre, B. Emert, C. Krepler, M. Beqiri, K. Sproesser, P. A. Brafford, M. Xiao, E. Eggan, I. N. Anastopoulos, C. A. Vargas-Garcia, A. Singh, K. L. Nathanson, M. Herlyn, A. Raj, Rare cell variability and drug-induced reprogramming as a mode of cancer drug resistance. *Nature* **546**, 431–435 (2017). doi:10.1038/nature22794 Medline

16. S. L. Spencer, S. Gaudet, J. G. Albeck, J. M. Burke, P. K. Sorger, Non-genetic origins of cell-to-cell variability in TRAIL-induced apoptosis. *Nature* **459**, 428–432 (2009). doi:10.1038/nature08012 Medline

17. M. Stoeckius, S. Zheng, B. Houck-Loomis, S. Hao, B. Z. Yeung, W. M. Mauck 3rd, P. Smibert, R. Satija, Cell Hashing with barcoded antibodies enables multiplexing and doublet detection for single cell genomics. *Genome Biol.* **19**, 224 (2018). doi:10.1186/s13059-018-1603-1 Medline

18. J. Gehring, J. H. Park, S. Chen, M. Thomson, L. Pachter, Highly multiplexed single-cell RNA-seq for defining cell population and transcriptional spaces. bioRxiv 315333 [Preprint] 5 May 2018. https://doi.org/10.1101/315333.

19. C. S. McGinnis, D. M. Patterson, J. Winkler, D. N. Conrad, M. Y. Hein, V. Srivastava, J. L. Hu, L. M. Murrow, J. S. Weissman, Z. Werb, E. D. Chow, Z. J. Gartner, MULTI-seq: Sample multiplexing for single-cell RNA sequencing using lipid-tagged indices. *Nat. Methods* **16**, 619–626 (2019). doi:10.1038/s41592-019-0433-8 Medline

20. D. Shin, W. Lee, J. H. Lee, D. Bang, Multiplexed single-cell RNA-seq via transient barcoding for simultaneous expression profiling of various drug perturbations. *Sci. Adv.* **5**, eaav2249 (2019). doi:10.1126/sciadv.aav2249 Medline

21. J. Cao, M. Spielmann, X. Qiu, X. Huang, D. M. Ibrahim, A. J. Hill, F. Zhang, S. Mundlos, L. Christiansen, F. J. Steemers, C. Trapnell, J. Shendure, The single-cell transcriptional landscape of mammalian organogenesis. *Nature* **566**, 496–502 (2019). doi:10.1038/s41586-019-0969-x Medline

22. M. A. McBrian, I. S. Behbahan, R. Ferrari, T. Su, T.-W. Huang, K. Li, C. S. Hong, H. R. Christofk, M. Vogelauer, D. B. Seligson, S. K. Kurdistani, Histone acetylation regulates

intracellular pH. *Mol. Cell* **49**, 310–321 (2013). doi:10.1016/j.molcel.2012.10.025 Medline

23. S. A. Comerford, Z. Huang, X. Du, Y. Wang, L. Cai, A. K. Witkiewicz, H. Walters, M. N. Tantawy, A. Fu, H. C. Manning, J. D. Horton, R. E. Hammer, S. L. McKnight, B. P. Tu, Acetate dependence of tumors. *Cell* **159**, 1591–1602 (2014). doi:10.1016/j.cell.2014.11.020 Medline

24. D. A. Cusanovich, R. Daza, A. Adey, H. A. Pliner, L. Christiansen, K. L. Gunderson, F. J. Steemers, C. Trapnell, J. Shendure, Multiplex single cell profiling of chromatin accessibility by combinatorial cellular indexing. *Science* **348**, 910–914 (2015). doi:10.1126/science.aab1601 Medline

25. J. Cao, J. S. Packer, V. Ramani, D. A. Cusanovich, C. Huynh, R. Daza, X. Qiu, C. Lee, S. N. Furlan, F. J. Steemers, A. Adey, R. H. Waterston, C. Trapnell, J. Shendure, Comprehensive single-cell transcriptional profiling of a multicellular organism. *Science* **357**, 661–667 (2017). doi:10.1126/science.aam8940 Medline

26. L. McInnes, J. Healy, UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction. arXiv:1802.03426 [stat.ML] (9 February 2018).

27. M. Jost, Y. Chen, L. A. Gilbert, M. A. Horlbeck, L. Krenning, G. Menchon, A. Rai, M. Y. Cho, J. J. Stern, A. E. Prota, M. Kampmann, A. Akhmanova, M. O. Steinmetz, M. E. Tanenbaum, J. S. Weissman, Combined CRISPRi/a-based chemical genetic screens reveal that rigosertib is a microtubule-destabilizing agent. *Mol. Cell* **68**, 210–223.e6 (2017). doi:10.1016/j.molcel.2017.09.012 Medline

28. G. Grosveld, T. Verwoerd, T. van Agthoven, A. de Klein, K. L. Ramachandran, N. Heisterkamp, K. Stam, J. Groffen, The chronic myelocytic cell line K562 contains a breakpoint in bcr and produces a chimeric bcr/c-abl transcript. *Mol. Cell. Biol.* **6**, 607–616 (1986). doi:10.1128/MCB.6.2.607 Medline

29. E. K. Greuber, P. Smith-Pearson, J. Wang, A. M. Pendergast, Role of ABL family kinases in cancer: From leukaemia to solid tumours. *Nat. Rev. Cancer* **13**, 559–571 (2013). doi:10.1038/nrc3563 Medline

30. J. Barretina, G. Caponigro, N. Stransky, K. Venkatesan, A. A. Margolin, S. Kim, C. J. Wilson, J. Lehár, G. V. Kryukov, D. Sonkin, A. Reddy, M. Liu, L. Murray, M. F. Berger, J. E. Monahan, P. Morais, J. Meltzer, A. Korejwa, J. Jané-Valbuena, F. A. Mapa, J. Thibault, E. Bric-Furlong, P. Raman, A. Shipway, I. H. Engels, J. Cheng, G. K. Yu, J. Yu, P. Aspesi Jr., M. de Silva, K. Jagtap, M. D. Jones, L. Wang, C. Hatton, E. Palescandolo, S. Gupta, S. Mahan, C. Sougnez, R. C. Onofrio, T. Liefeld, L. MacConaill, W. Winckler, M. Reich, N. Li, J. P. Mesirov, S. B. Gabriel, G. Getz, K. Ardlie, V. Chan, V. E. Myer, B. L. Weber, J. Porter, M. Warmuth, P. Finan, J. L. Harris, M. Meyerson, T. R. Golub, M. P. Morrissey, W. R. Sellers, R. Schlegel, L. A. Garraway, The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. *Nature* **483**, 603–607 (2012). doi:10.1038/nature11003 Medline

31. C. Dai, S. Santagata, Z. Tang, J. Shi, J. Cao, H. Kwon, R. T. Bronson, L. Whitesell, S. Lindquist, Loss of tumor suppressor NF1 activates HSF1 to promote carcinogenesis. *J. Clin. Invest.* **122**, 3742–3754 (2012). doi:10.1172/JCI62727 Medline

32. L. Haghverdi, A. T. L. Lun, M. D. Morgan, J. C. Marioni, Batch effects in single-cell RNA-sequencing data are corrected by matching mutual nearest neighbors. *Nat. Biotechnol.* **36**, 421–427 (2018). doi:10.1038/nbt.4091 Medline

33. C. Trapnell, D. Cacchiarelli, J. Grimsby, P. Pokharel, S. Li, M. Morse, N. J. Lennon, K. J. Livak, T. S. Mikkelsen, J. L. Rinn, The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. *Nat. Biotechnol.* **32**, 381–386 (2014). doi:10.1038/nbt.2859 Medline

34. W. Brazelle, J. M. Kreahling, J. Gemmer, Y. Ma, W. D. Cress, E. Haura, S. Altiok, Histone deacetylase inhibitors downregulate checkpoint kinase 1 expression to induce cell death in non-small cell lung cancer cells. *PLOS ONE* **5**, e14335 (2010). doi:10.1371/journal.pone.0014335 Medline

35. J.-S. Roe, F. Mercan, K. Rivera, D. J. Pappin, C. R. Vakoc, BET bromodomain inhibition suppresses the function of hematopoietic transcription factors in acute myeloid leukemia. *Mol. Cell* **58**, 1028–1039 (2015). doi:10.1016/j.molcel.2015.04.011 Medline

36. J. E. Brownell, J. Zhou, T. Ranalli, R. Kobayashi, D. G. Edmondson, S. Y. Roth, C. D. Allis, Tetrahymena histone acetyltransferase A: A homolog to yeast Gcn5p linking histone acetylation to gene activation. *Cell* **84**, 843–851 (1996). doi:10.1016/S0092-8674(00)81063-6 Medline

37. J. Taunton, C. A. Hassig, S. L. Schreiber, A mammalian histone deacetylase related to the yeast transcriptional regulator Rpd3p. *Science* **272**, 408–411 (1996). doi:10.1126/science.272.5260.408 Medline

38. S. K. Kurdistani, Chromatin: A capacitor of acetate for integrated regulation of gene expression and cell physiology. *Curr. Opin. Genet. Dev.* **26**, 53–58 (2014). doi:10.1016/j.gde.2014.06.002 Medline

39. K. E. Wellen, G. Hatzivassiliou, U. M. Sachdeva, T. V. Bui, J. R. Cross, C. B. Thompson, ATP-citrate lyase links cellular metabolism to histone acetylation. *Science* **324**, 1076–1080 (2009). doi:10.1126/science.1164097 Medline

40. A. R. Quinlan, I. M. Hall, BEDTools: A flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842 (2010). doi:10.1093/bioinformatics/btq033 Medline

41. C. Ritz, F. Baty, J. C. Streibig, D. Gerhard, Dose-response analysis using R. *PLOS ONE* **10**, e0146021 (2015). doi:10.1371/journal.pone.0146021 Medline

42. M. E. Ritchie, B. Phipson, D. Wu, Y. Hu, C. W. Law, W. Shi, G. K. Smyth, limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* **43**, e47 (2015). doi:10.1093/nar/gkv007 Medline

43. I. Tirosh, B. Izar, S. M. Prakadan, M. H. Wadsworth 2nd, D. Treacy, J. J. Trombetta, A. Rotem, C. Rodman, C. Lian, G. Murphy, M. Fallahi-Sichani, K. Dutton-Regester, J.-R. Lin, O. Cohen, P. Shah, D. Lu, A. S. Genshaft, T. K. Hughes, C. G. K. Ziegler, S. W. Kazer, A. Gaillard, K. E. Kolb, A.-C. Villani, C. M. Johannessen, A. Y. Andreev, E. M. Van Allen, M. Bertagnolli, P. K. Sorger, R. J. Sullivan, K. T. Flaherty, D. T. Frederick, J. Jané-Valbuena, C. H. Yoon, O. Rozenblatt-Rosen, A. K. Shalek, A. Regev, L. A. Garraway, Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. *Science* **352**, 189–196 (2016). doi:10.1126/science.aad0501 Medline

44. L. Väremo, J. Nielsen, I. Nookaew, Enriching the gene set analysis of genome-wide data by incorporating directionality of gene expression and combining statistical hypotheses and methods. *Nucleic Acids Res.* **41**, 4378–4391 (2013). doi:10.1093/nar/gkt111 Medline

45. A. Subramanian, P. Tamayo, V. K. Mootha, S. Mukherjee, B. L. Ebert, M. A. Gillette, A. Paulovich, S. L. Pomeroy, T. R. Golub, E. S. Lander, J. P. Mesirov, Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. U.S.A.* **102**, 15545–15550 (2005). doi:10.1073/pnas.0506580102 Medline