Disrupting cellular memory to overcome drug resistance

Guillaume Harmange[1], Raúl A. Reyes Hueros[3], Dylan Schaff[2], Benjamin Emert[5], Michael Saint-Antoine[4], Shivani Nellore, Mitchell E. Fane[6], Gretchen M. Alicea[6], Ashani T. Weeraratna[6,7], Abhyudai Singh[4], Sydney M. Shaffer[2,8]

[1] Cellular and Molecular Biology Graduate Group, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA

[2] Department of Bioengineering, School of Engineering and Applied Sciences, University of Pennsylvania, Philadelphia, PA, USA

[3] Department of Biochemistry and Molecular Biophysics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA

[4] Department of Electrical and Computer Engineering, University of Delaware, Newark, DE 19716, USA

[5] Division of Biology and Biological Engineering, California Institute of Technology, Pasadena, CA 91125, USA

[6] Department of Biochemistry and Molecular Biology, Johns Hopkins School of Public Health, Baltimore, MD, USA.

[7] Sidney Kimmel Cancer Center, Johns Hopkins School of Medicine, Baltimore, MD, USA.

[8] Department of Pathology, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA.

Abstract

Cellular memory describes the length of time a particular transcriptional state exists at the single-cell level. Transcriptional states with memory can underlie important processes in biology, including therapy resistance in cancer. Here we present a new experimental and computational approach for identifying gene expression states with memory at single-cell resolution by combining single-cell RNA sequencing (scRNA-seq) with cellular barcoding. With this technique, we can systematically quantify the full expression profile of these memory states as well as their population dynamics including the relative growth rates of cells within different states and the rates of switching between states. We applied this approach to human melanoma cells and uncovered memory gene expression states that are predictive of which cells will be resistant to combination BRAF and MEK inhibition. While these cells have not been treated with targeted therapies, they already express markers of resistance, and thus are referred to here as “primed” for resistance. From the scRNA-seq data alone, the drug-susceptible and primed cells appear as two distinct and unrelated cell populations. From the lineage barcodes, we found that most cells remain within the same state, demonstrating memory of gene expression. However, we also directly observe state switching as we find that about 18% of lineages contain cells that have

switched states. While the molecular drivers of state switching are not immediately apparent from scRNA-seq data alone, we specifically analyzed lineages that undergo state switching and identified TGF-β and PI3K as drivers of state switching at the single-cell level. Through functional validation and single-cell RNA-sequencing, we show mechanistic single-cell manipulation of state switching that ultimately yields changes in cellular phenotype. We leverage these mechanisms of state switching to delay resistance by demonstrating that reducing the number of primed cells with a PI3K inhibitor prior to BRAF and MEK inhibition ultimately leads to fewer resistant colonies. Our results show that modulation of cellular signaling can globally regulate gene expression programs to overcome therapy resistance.

Introduction

Gene expression memory describes the length of time that a particular expression state exists in an individual cell or lineage of cells. Memory of gene expression states can exist over a range of different timescales and is ultimately a quantitative measurement (Angel et al. 2011; Shaffer et al. 2020). For cancer, an intermediate timescale of memory has been associated with a number of important phenotypes including, stemness and differentiation (Chaligne et al. 2021; Gupta et al. 2011), metastasis (Kaur et al. 2022), and therapy resistance (Shaffer et al. 2020; Torre et al. 2021; Sharma et al. 2010; Oren et al. 2021). In these examples, the cellular state underlying the phenotype is stable enough to persist through multiple cell divisions, but is ultimately not permanent, and thus amenable to switching states. For cancer therapy resistance, drug-naive cells seemingly randomly fluctuate between two states, one which is susceptible to targeted therapy and another which would become resistant if the drug is applied, termed “primed” for drug resistance (Emert et al. 2021; Shaffer et al. 2017). However, the molecular drivers underlying these heritable, but reversible memory states remain unknown.

Because drug resistance can emerge from cells that are in a primed gene expression state, one intriguing possibility to prevent resistance is to transform primed cells into drug-susceptible cells. However, because these gene expression states have memory and are heritable through cell division, we need to identify the molecular cues that enable cells to switch between these states. Knowing these cues would make it possible to therapeutically target these pathways to drive cells out of the primed gene expression state to overcome drug resistance and provide mechanistic insight into how state switching occurs. Such an approach requires deep characterization of the processes underlying state switching, which remains exceedingly difficult with current methods.

Currently, there are limited techniques that will reveal memory and state switching in single cells. Our previous work inferred cellular memory from bulk RNA-seq measurements (Shaffer et al. 2020), but failed to capture drivers of switching between memory states. Meanwhile, scRNA-seq can capture the heterogeneity of a population, but it fails to capture the timescales for which different states have been present in individual cells, and lacks the resolution to capture transitioning gene expression states. A number of computational and experimental techniques have been developed to resolve time in single-cells on short timescales, on the order of hours (La Manno et al. 2018; Qiu et al. 2020), but few exist for the longer timescales of days to weeks as needed to track cellular memory. Recent advances in high-throughput cellular barcoding technologies have made it possible to track cellular lineages across any length of time (Bhang et

al. 2015). Pairing cellular barcoding with scRNA-seq enables us to now match cellular lineages with their transcriptome (Biddy et al. 2018; Oren et al. 2021; Weinreb et al. 2020; Emert et al. 2021), and thus is an ideal tool for tracking transcriptional states across lineages to measure cellular memory.

Here, we leveraged recent advances in high-throughput cellular barcoding and scRNA-seq to develop scMemorySeq, an experimental and computational framework for studying reversible cellular states with memory. We applied scMemorySeq to drug naive melanoma cells which are known to fluctuate between the drug-susceptible and primed state (Emert et al. 2021; Shaffer et al. 2017). We found that the drug-susceptible and primed states are two very distinct, high-memory states that appear unrelated based only on scRNA-seq alone; however, through cellular barcoding, we identified lineages containing cells from both states, directly showing that single cells can switch between states. While the molecular drivers of state switching were not immediately apparent from scRNA-seq alone, by analyzing lineages that lose memory of their state, we identified TGF-β and PI3K as mechanisms of modulating state switching. Ultimately, we found that by initially disrupting the primed state through PI3K inhibition and then applying BRAF inhibitor in combination with a MEK inhibitor (BRAFi/MEKi) to kill the cells, we could reduce the frequency of therapy resistance. Taken together, we demonstrate the feasibility of molecularly targeting state switching to eliminate gene expression states in cancer that prime cells for undesirable phenotypes.

Results

We first sought to identify the molecular pathways underlying cellular memory in the drug-susceptible and primed cells (Fig 1A). Towards this goal, we developed a new approach called scMemorySeq which uses cellular barcoding combined with scRNA-seq to measure cellular memory. In our experimental design, the cellular barcodes enable high-throughput tracking of cells over any desired period of time and the scRNA-seq reveals the transcriptional states of every cell. Ultimately, we infer cellular memory by examining gene expression across individual lineages. For genes or sets of genes that have memory over the experimental timescale, all cells from the same lineage should be in the same end state and match the initial state (Fig 1B). However, if cellular states change over the experimental timescale, then the end-state will contain cells with multiple gene expression states, some of which are different from the initial state. We applied this approach to BRAF-mutant melanoma cells in which a rare subpopulation of cells are resistant to BRAF inhibitors. We transduced WM989 melanoma cells with a high-complexity viral barcode library consisting of a transcribed 100 base pair semi-random barcode sequence in the 3' UTR of GFP (Emert et al. 2021) (Fig. 1C). In this experimental model, our phenotype of interest (cells that are primed for resistance) is rare. To improve our ability to capture primed cells, and to establish the initial state of lineages, we sorted cells based on known primed cell markers *(*EGFR+ and NGFR+) and also isolated a matched control population (Fig. 1C). To measure cellular memory, we allowed the cells to expand through roughly 5 doublings (12-14 days). Over this time, we expect that most cells will maintain memory of their initial gene expression state, but that a subset of cells will lose memory of their initial state, thereby capturing the reversion of these expression states. At the endpoint, we harvested these cells for scRNA-seq and performed a PCR side reaction to specifically amplify the lineage barcodes from each cell.

We profiled a total of 12,531 (7,581 cells with barcodes) melanoma cells and found two major transcriptionally distinct populations (Fig 1D, Supp. Fig. 1A). We found that one cluster showed much higher expression of many of the genes associated with the primed resistant state including *EGFR* and *AXL* (Shaffer et al. 2020; Emert et al. 2021; Shaffer et al. 2017), and that most of the sorted NGFR and EGFR-high cells landed in this cluster (Fig. 1D, Supp. Fig. 1B). Of note, the two clusters were sufficiently distinct that they clustered separately even when the NGFR and EGFR enriched sample was not included in the analysis (Supp. Fig. 1D). Consistent with this observation, the other cluster showed higher expression of genes associated with the drug-susceptible state including *SOX10* and *MITF* (Fig. 1D Supp. Fig. 1E). Within the cluster expressing primed cell genes, we noticed significant heterogeneity across previously described marker genes including *EGFR*, *AXL* and *NGFR* with each gene expressed by some of the cells in the cluster, but not all of the cells (Fig. 1D and Supp. Fig. 1E). In a separate experiment, we profiled the chromatin accessibility of EGFR and NGFR-high cells, and found epigenetic differences between these populations further confirming that there is additional heterogeneity within the primed state (Supp. Fig. 1 F). Thus, we suspected that our markers from previous work (Shaffer et al. 2017) did not capture all of the cells within this cluster.

In order to capture all of the primed cells, we identified a novel marker, *NT5E*, that encapsulated the entire cluster containing the primed population (Fig. 1D). To validate *NT5E* as an effective marker of primed cells, we labeled WM989 cells with NT5E antibody and sorted the top 2% of cells. We then applied the targeted BRAF inhibitor, vemurafenib, and found that the NT5E samples had 5.5 fold more resistant cells compared to a mix control (23,305 cells in NT5E-high and 4,249 cells in mixed condition) (Fig 1E, Supp. Fig. 1G). Thus, we demonstrated that the marker gene NT5E captures the entire cluster of transcriptionally similar cells and that these cells are indeed more resistant to targeted therapy. Surprisingly, despite the fact that our previous work showed that cells can switch between the drug-susceptible and primed state (Shaffer et al. 2020), there are no cells visibly transitioning between these populations in UMAP space. The lack of transitioning cells was also observed in a replicate WM989 scRNA-seq experiment (Supp. Fig. 2A). This highlights the fact that looking at the scRNA-seq data alone is not always enough to uncover cellular dynamics and state switching.

To measure the frequency of primed cells in vivo, we used RNA FISH HCRv3.0 on tumor samples derived from WM989 cells grown in NOD/SCID mice (Torre et al. 2021). We used probes for *NT5E*, and *SOX10* on tumor sections and quantified expression across 5,600 cells. We found rare cells expressing high levels of *NT5E* mRNA scattered throughout the tissue (Fig. 1F, additional images in Supp. Fig. 1H). *SOX10* showed diffuse expression across the tissue, but many of the *NT5E*-high cells did not have *SOX10* expression, as predicted by the scRNA-seq. Altogether these *NT5E*-high *SOX10*-low cells demonstrate that the primed state exists both in vitro and in vivo (and is not an artifact of cell culture conditions). We also performed scRNA-seq on a different melanoma cell line, WM983B, and found a subpopulation population of cells with a large number of primed cell markers, including *NT5E* (Supp. Fig. 2B). To establish whether these cell states exist in patients tumors, we analyzed published scRNA-seq data (Tirosh et al. 2016; Jerby-Arnon et al. 2018) which included 7 samples directly from human tumors and found that 5

of them had a subpopulation of cells with high expression of genes associated with the primed state (including the EMT and TGF-β gene signatures, Supp. Fig. 2C). Overall these findings suggest that our WM989 cell line serves as a representative model for gene expression states that exist in patients.

With the scRNA-seq data alone, it is not possible to know whether cells are rapidly fluctuating between these states or whether the states persist at the single-cell level through multiple cell divisions. To extract the time that cells remain in these states, we turned to our barcoding data. First, we used the scRNA-seq data to assign all cells to one of two states, either drug-susceptible or primed for drug resistance. We then classified each lineage into one of four categories (Fig 2A and 2B): 1. Drug-susceptible lineage (containing only drug-susceptible cells), 2. Primed lineage (containing only cells in the primed state), 3. Switching from drug-susceptible to the primed state (initial state is drug-susceptible and lineage contains drug-susceptible and primed cells), or 4. Switching from primed to drug-susceptible state (initial state is primed and lineage contains drug-susceptible and primed cells). Across the different classes of lineages, we found that the lineage sizes were largest for the entirely drug-susceptible lineages and smallest for the primed lineages, which is consistent with scRNA-seq predictions of cell cycle differences between the two clusters (Fig. 2A Supp. Fig. 1C). Given the high memory observed in both the drug-susceptible and primed states, we compared the differentially expressed genes between these states captured by scMemorySeq to bulk measurements of gene expression memory performed in the same cell line from (Shaffer et al. 2020). We observed a strong correlation between the two methods on a per gene basis (Supp. Fig. 2F). To derive the rates of proliferation and state switching from this data, we used a stochastic two-state model. Using the model we estimate that primed cells proliferate at approximately half the rate of drug-susceptible cells on average. We also estimate that drug-susceptible cells switch to the primed state once every 135-233 cell divisions, while primed cells are estimated to switch to the drug-susceptible state once every 5-8 cell divisions. The large difference in switching rates is due to the rarity of the primed cell state which must have a faster switching rate to maintain a constant proportion of primed cells at steady state. Thus, on the single-cell level, the drug-susceptible state is significantly more stable than the primed state (See Supp. Method 1).

After classifying the lineages based on whether they switch between states, we next wondered if there were transcriptional differences between lineages that undergo state-switching and lineages that do not switch. We hypothesized that these differences might exist if there was an intermediate transcriptional state that cells pass through as they transition between the drug-susceptible and the drug-primed states. To uncover such a state, we compared lineages that contain only drug-susceptible cells to lineages that switch from drug-susceptible to the primed state (Fig. 2C). Importantly, this comparison requires the assumption that intermediate states would also demonstrate memory through cell division (schematic in Supp. Fig. 2D). Consistent with the presence of such an intermediate state, we found 575 genes differentially expressed between these types of lineages (based upon a differential gene expression analysis with a cutoff of 0.05 Bonferroni adjusted p-value, and a 0.25 log fold change cutoff) (Fig. 2D, Supp. Table 1). We then used this gene list to develop a “crossing-lineage score”, applied this score to all cells in the data set, and then projected them into UMAP space (Fig. 2E). We identified two primary regions within

the drug-susceptible cluster in UMAP space that show enrichment for cells in this crossing expression state. We then set a threshold score and classified cells as “intermediate cells” in addition to the two previously defined states, drug-susceptible and drug-primed (Fig. 2F). These intermediate cells are very rare, found at a frequency of around 2% in the population and would have not been identified using standard scRNA-seq techniques and analyses.

We next wanted to know what pathways are activated in the intermediate cell state. We first used gene set enrichment analysis on the differentially expressed gene list from Fig. 2D. The top pathways included UV response down, epithelial-to-mesenchymal transition (EMT), and response to hypoxia (Fig. 2H). We then examined the activity of these different pathways in UMAP space and noted that the EMT pathway score localized in the same regions of the UMAP as the intermediate cell population (Fig. 2I). We focused on this gene set as multiple lines of evidence in melanoma suggest that an “EMT-like” gene expression state is associated with resistance to BRAFi/MEKi (Ramsdale et al. 2015; Rambow et al. 2018; Verfaillie et al. 2015; Tang et al. 2020; Pedri et al. 2022). Another enriched pathway in our analysis was TGF-β signaling, which also showed enrichment in the same region of the UMAP as the intermediate cells (Fig. 2J). Furthermore, both the EMT and TGF-β pathways were further enriched in the primed cells beyond the levels in intermediate cells, suggesting that these are early changes as cells transition from drug-susceptible to intermediate to drug-primed (Fig. 2I and 2J). We also observed that these intermediate cells maintain expression of many genes associated with the drug-susceptible state, including melanocyte identity genes *SOX10*, and *MITF*, but have already begun to upregulate some of the important primed cell marker genes including *FN1* and *SERPINE2* (Fig 2G). A few genes were uniquely expressed only in the intermediate state cells including *NFATC2*, and *MGP*. Intriguingly, *NFATC2* was previously found to be a regulator of *MITF* and melanoma dedifferentiation (Perotti et al. 2016; Aibar et al. 2017; Shaffer et al. 2020).

Given the role of TGF-β as a potent inducer of EMT and the enrichment for the TGF-β pathway in the intermediate cells, we hypothesized that applying TGFB1 to drug-susceptible cells might drive them into the primed state (Fig. 3A) (Xu, Lamouille, and Derynck 2009). Importantly, we also observed that TGFB1 and its receptor are highly upregulated in primed cells suggesting that these cells naturally activate TGF-β signaling through autocrine or paracrine mechanisms (Supp. Fig. 2E). To test if TGFB1 induces the primed state, we treated melanoma cell lines including WM989 and WM983B with recombinant TGFB1 for 5 days and then performed flow cytometry for primed cell marker gene NT5E (Fig. 3B). We found that treatment with TGFB1 increased the percentage of cells of primed cells in WM989 from 1.98% to 19.15% and in WM983B from 10.04% to 81.56% (Fig. 3C and 3D). Our observed effects of TGFB1 on melanoma are in agreement with literature showing that TGF-β signaling can induce a dedifferentiation state (Lee et al. 2020; Sun et al. 2014). The dedifferentiated state is similar to our primed cell population as these cells are also characterized by their decreased expression of melanocyte transcription factors SOX10 and MITF (Tirosh et al. 2016; Sun et al. 2014; Lee et al. 2020; Hoek et al. 2008). Furthermore, our work shows that these dedifferentiation events occur in a rare subpopulation of melanoma cells before targeted therapy whereas other studies focus on these gene expression states after treatment with targeted therapy.

Since TGFB1 was sufficient to induce the primed state, we wondered whether inhibiting TGF-β signaling would force cells in the opposite direction into the drug-susceptible state. Thus, we treated WM989 cells with LY2109761, a targeted TGFBR1/2 inhibitor (TGFBRi), for 5 days. Unexpectedly, we found that the TGFBRi did not change the percentage of primed cells significantly (Fig. 3C). To confirm the specificity of the inhibitor, we treated WM989 cells with both recombinant TGFB1 and TGFBRi and found that indeed the inhibitor was able to block the effects of TGFB1 (Fig. 3C). Overall this suggests that TGF-β signaling is sufficient to increase the percentage of primed cells, but is not necessary for maintaining them in the population.

Because the TGFBRi was not able to drive cells out of the primed state, we wondered whether other signaling pathways might be involved in maintaining the primed state. We noted that a number of growth factors including *FGF1*, *VGF*, *BDNF*, *VEGFA*, *VEGFC*, and *PDGFA* were all upregulated in primed cells. Many of these growth factors have downstream PI3K activation as a common pathway (Stommel et al. 2007). We also noted that while the canonical TGF-β signaling has downstream SMAD family effectors (Derynck and Zhang 2003), non-canonical TGF-β signaling can activate PI3K (Bakin, Tomlinson, and Bhowmick, n.d.; Zhang 2009). Given these common mechanisms potentially converging on PI3K, we wondered whether blocking PI3K could eliminate the primed state through blocking both the effects of TGF-β signaling and other upregulated growth factors and receptors. We tested the effects of blocking PI3K on the primed state by treating WM989 and WM983B cells with PI3K inhibitor (PI3Ki), GDC-0941, for 5 days and then performing flow cytometry for NT5E to quantify the percentage of primed cells. We found that the PI3Ki decreased the percentage of primed cells from 1.98% to 0.31% and 10.04% to 3.66% in WM989 and WM983B, respectively (Fig 3C and 3D). Furthermore, we tested whether the PI3Ki is able to block the effects of TGF-β signaling by simultaneously treating WM989 cells with both TGFB1 and the PI3Ki. Indeed, we found that the PI3Ki was able to block the increase in primed cells seen when we treat with TGFB1 alone (Fig 3C), suggesting that the TGFB1-mediated effects on priming melanoma cells acts through downstream PI3K activity.

Given that numerous growth factors were differentially expressed between the drug-susceptible and primed for resistance states, we wondered whether other ligands, in addition to TGFB1, might also be able to induce state switching. We treated WM989 cells with EGF, BDNF, and IL6 for 5 days (Supp. Fig. 3D) and found that none of these factors were able to induce the primed cell state as seen with TGFB1 despite their potential to increase PI3K signaling (Simiczyjew et al. 2019; Meng et al. 2019; Wegiel et al. 2007). Thus, we concluded that TGFB1 is a unique inducer of state switching and other growth factors do not have the same functionality.

Since TGFB1 and the PI3Ki were both able to change the percentage of cells in the primed state, we wondered how this is achieved at the single cell level. Specifically, does TGFB1 force more cells to switch into the primed state? Conversely, does the PI3Ki force cells to exit the primed state? Based on flow cytometry alone, we can conclude that the percentage of the population is shifted by these treatments; however, changes in multiple different parameters could lead to this same effect. For instance, the result that TGFB1 increases the percentage of primed cells could be explained by (1) an increase in growth rate of primed cells, (2) a selective killing of drug-susceptible cells, or (3) state-switching from the drug-susceptible into the drug-primed state. The

opposite consideration is necessary with the finding that the PI3Ki decreases the percentage of primed cells. This result could be the effects of (1) an increase in growth rate in the drug-susceptible cells, (2) a selective killing of primed cells, or (3) state-switching from the primed into the drug-susceptible state. Thus, to directly test how TGFB1 and PI3K shift the percentage of primed cells, we redesigned the scMemorySeq approach leveraging the knowledge of memory to test the effects of these perturbations. For this experimental design, we transduced WM989 cells with the high-complexity transcribed barcode library and then allowed the cells to go through 7-8 doublings based on experiments demonstrating that memory of these states still persists through these timescales (Supp. Fig. 3A-3C). We then split the cells into 4 separate plates and subjected each to a different condition (untreated, TGFB1 recombinant protein, the TGFBRi, and the PI3Ki) (Fig. 4A). Since cells within the same lineages have very similar gene expression states due to memory, this allows us to approximate studying how the “same cell” will react to the different conditions. After 5 days, we harvested each sample and performed scRNA-seq and barcode-sequencing to capture the transcriptional state and barcode under each condition (Fig. 4A). With this experimental design, each barcode is represented across all the conditions, and thus, we can see the effects of that treatment by comparing the cells in each treatment to the cells in the untreated plate with the same barcode.

We analyzed a total of 40,021 cells across the four different samples and found similar transcriptional states distributed across UMAP space as in previous experiments (Fig. 4B, Supp. Fig. 4B) with *SOX10*-low cells clustered separately from cells with high expression of *EGFR*, *AXL*, and *NT5E*. Across the different treatment conditions, we found that the TGFB1 treatment shifted cells towards the primed state, while the PI3Ki shifted cells towards the drug-susceptible state (Fig. 4C). Consistent with previous experiments, the TGFBRi had minimal effects on the primed cell population. To classify cells as primed vs. drug-susceptible, we selected clusters 4 and 10 from the high-dimensional clustering in Seurat, as these clusters contained the majority of cells expressing the known primed cell marker genes (Fig. 4D, Supp. Fig. 4A). Based on this classification, we quantified the percentage of primed cells based upon the full transcriptional state captured by scRNA-seq and further confirmed the effects of TGFB1 and the PI3Ki (Fig. 4E). Importantly, the primed cell state induced by treatment with TGFB1 was transcriptionally very similar to untreated cells in the primed state (Fig. 4F). To explicitly test for differences, we performed differential gene expression to compare the TGFB1-induced and untreated primed cells. We found that some genes, including *NGFR*, *FGFR1*, *FOSL1*, and *JUN*, were induced by the TGFB1 treatment, but were not as highly expressed as in the untreated primed cells. This might be significant as *NGFR* expression has been linked to invasive properties of melanoma cells (Radke, Roßner, and Redmer 2017; Filipp, Li, and Boiko 2019). We also noted that TGFB1 seemed to induce even higher expression for many of the primed cell genes including *FN1*, *SERPINE1*, *COL1A1* and *VGF* (Supp. Fig. 4D). This might be explained by the high dose of TGFB1 used in these experiments relative to what these cells would generate on their own, which could lead to a more dramatic effect on the cells.

Next, we analyzed the barcoding data to test for state switching at the single cell level. Our data set included 19,740 lineages (containing a minimum of 3 cells per lineage) with 49% of lineages represented across all four conditions. To test whether TGFB1 induces more primed cells, we first

evaluated lineages in which the control sample had cells exclusively in the drug-susceptible state. Across each of these lineages, we quantified the percentage of primed cells in the other conditions and found that the matched set of lineages in the TGFB1-treated condition had a higher fraction of cells in the primed state (Fig. 5A). Reassuringly, the TGFBRi and the PI3Ki only had minor increases in the percentage of primed cells across these lineages. We next tested for the opposite direction, switching from the primed state into the drug-susceptible state. We analyzed lineages in which the entire lineage was in the primed state in the untreated sample (of note, since the primed state is rare, we only had 8 lineages that met this criteria). In the PI3K-treated sample, the same lineages switched states and ultimately had a lower percentage of primed cells (Fig. 5B). We noted that while 6 of the 8 lineages switched out of the primed state in the PI3Ki, 2 lineages did not respond to the PI3Ki. It is possible that non-responsive lineages would require a longer treatment with the PI3Ki to fully shift into the drug-sensitive gene expression state.

To extend our analysis to include all of the lineage data, we developed a stochastic model of state switching between the drug-susceptible and drug-primed state. Our model included different state switching parameters (k_{on}, k_{off}), different growth rates, and different death rates in each cell state (Fig. 5C). We used the model to simulate experimental data for different scenarios in which different parameters are changing. For instance, for TGFB, the increase in primed cells could be explained by 3 possible parameter changes, 1) increase in K_{on} for the primed state, 2) increase in primed cell proliferation rate, or 3) increase in death rate among drug-susceptible cells. To constrain the proliferation rate parameter in our model, we performed live-cell imaging to measure the direct effects of 5 days of TGFB1 and the PI3Ki on the proliferation rates of primed and drug-susceptible cells. Across conditions, we found that the primed cells proliferate more slowly than the drug-susceptible cells (Supp. Fig. 5B). Additionally, we found that treatment with either TGFB1 or the PI3Ki decreased growth rate by similar amounts in each population (Sup. Fig. 5B).

Given these constraints on proliferation rates, we then ran 1 million simulations of the model varying each parameter and found that increasing the rate at which the cells switch into the primed state is the best fit for our data. Interestingly, the model also suggested that drug-susceptible cells in lineages with a high proportion of cells already in the primed state are more easily switched into the primed state (See Supp. Method 1). We similarly considered possible parameter changes that could account for the decrease in primed cells upon PI3Ki treatment. Given experimentally measured constraints on proliferation rates, we found that an increased rate of cells switching from the primed state to the drug-susceptible state was the best fit for our data (See Supp. Method 1). Taken together, this data shows that TGFB1 and the PI3Ki can induce state switching and that the observed changes in the number of primed cells are not due to other population dynamics.

Based upon the finding that TGFB1 induced more cells to enter the primed state, we next tested whether induced primed cells are more resistant to therapy. Our experimental design tested a sequential dosing strategy in which we first pretreat cells with TGFB1 and then BRAFi/MEKi to test for subsequent resistance (Fig. 6A). Unexpectedly, we found that, although fewer total cells died, pretreatment with TGFB1 led to accelerated killing of the remaining drug-sensitive cells following BRAFi/MEKi compared to cells treated with BRAFi/MEKi alone (Supp. Fig. 5C).

Furthermore, after 4 weeks of BRAFi/MEKi, we found that the pretreated TGFB1 sample had fewer resistant cells than the BRAFi/MEKi only sample. Interestingly, although there were fewer cells on the plate, they were much more evenly distributed suggesting that more of the initial cells survived BRAFi/MEKi, but that these cells were slow to proliferate (Fig 6B, Supp. Fig. 5C).

Given the result with the PI3Ki, driving cells out of the primed state, we wondered whether the PI3Ki could be combined with targeted therapy to reduce resistance. We used the same pretreatment experiment design as with TGFB1 in which we first treat with the agent to disrupt the primed state (here the PI3Ki) and then apply the targeted therapy (BRAFi/MEKi) to test for subsequent resistance (Fig. 6A). We found that pretreatment for 5 days with the PI3Ki followed by BRAFi/MEKi for 3 weeks decreased both the number of resistance colonies by 38% and the number of resistant cells 43% compared to BRAFi/MEKi alone (Fig. 6B representative images, Fig. 6C quantification across 6 wells and 3 replicates). We also performed the same experiment using live-cell imaging and acquiring images every 12 hours (Sup Fig. 5C, Supp. Video 1, Supp. Video 2). We noticed that while resistance does eventually develop in the sample pretreated with the PI3Ki, it developed later than in the BRAF/MEK inhibitor alone, demonstrating that pretreatment with the PI3Ki delayed the onset of resistance.

Discussion

Here we show that scMemorySeq is a powerful method for tracking gene expression memory in single cells. Our approach leverages the combination of scRNA-seq and cellular lineage barcoding to quantify memory of gene expression states in single cell data. When applied to melanoma cells that switch between two gene expression states, a drug-susceptible state and a state primed for drug resistance, we directly observed memory of these expression states in most cells and rare events in which cells switched between states. Further, we leveraged this data to find molecular drivers of state switching. By analyzing the gene expression differences in lineages that switch states, we identified and then tested TGFB and PI3K as mediators of state switching at the single-cell level. Ultimately, we show that manipulating state switching can alter cell phenotypes, and for the PI3Ki specifically, it can reduce drug resistance.

Broadly, it is intriguing that modulating signaling alone is sufficient to globally modify gene expression states in single cells and to affect their susceptibility to drugs. Here, PI3Ki is driving cells into a MAPK dependent transcriptional state, sensitizing cells to MAPK inhibitors. Importantly, a number of papers have reported the use of PI3Ki to reduce resistance in melanoma (Villanueva et al. 2010; Irvine et al. 2018; Mendoza, Er, and Blenis 2011; Nymark Aasen et al. 2019). Distinguishing our use of a PI3Ki, here, we find that the PI3Ki can reduce drug resistance even when only applied briefly before the addition of targeted therapy.This approach may lay out a generalizable strategy for reducing drug resistance in which perturbing signaling can globally tune gene expression to achieve susceptibility to drugs. In such a strategy, modulating signaling pathways would be used prior to the addition of the main targeted therapy to drive heterogeneous populations of cells into a drug susceptible state. This is in contrast to dosing with a combination of inhibitors at the same time, which is not always tolerated due to toxicity and side effects (Park et al. 2013; Rafsanjani Nejad et al. 2021; Jardim et al. 2020).

Conceptually, the idea of leveraging the plasticity of cells to revert them into a drug-susceptible state to delay drug resistance has previously been described in the context of drug holidays (Kavran et al. 2022). In contrast to our approach, in which we actively drive cells into a drug-susceptible state, a drug holiday is a break from the drug to alleviate the selective pressure thus giving the drug-resistant cells the opportunity to switch back to a drug-susceptible state at their intrinsic rate. Given the previous success of drug holidays (Das Thakur et al. 2013; Kavran et al. 2022), one strategy that we believe should be further explored is to use state switching drugs during what would be the holiday period. With this approach, the switch to a drug-susceptible state would be accelerated during the drug holiday, leading to potentially even less resistance. Future work is still needed to model these different scenarios and to experimentally test the efficacy of such dosing strategies.

Extending scMemorySeq beyond state switching in melanoma, there are a number of biological contexts in which specific drugs or ligands could be shifting gene expression states and population dynamics simultaneously. This is particularly relevant in cancer where there is a growing body of literature describing considerable heterogeneity at the single-cell level that shows variable degrees of memory (Tirosh et al. 2016; Barkley et al. 2021; Neftel et al. 2019). Here, we show that the scMemorySeq approach can provide a detailed systematic overview of these populations, and can also be used to test how closely related cells from the same lineage will react to different conditions. We believe that this approach is generalizable and could be used in other contexts to profile cell state transitions under different drugs, ligands, or environmental conditions. This is a growing area of interest as multiple recent studies showed that microenvironment and growth conditions of cancer cells can globally change both gene expression and sensitivity to drugs (Raghavan et al. 2021; Neal et al. 2018; Neftel et al. 2019).

In sum, we show how scMemorySeq uses cellular barcoding to reveal single-cell dynamics of drug resistance in melanoma. By tracking the memory of gene expression states we can identify stable cell populations as well as the factors that cause cells to change states. This approach can be widely applied to discover unknown dynamics in heterogeneous cell populations and to identify the key factors responsible for gene expression state changes in biological systems.

Methods

Antibodies

NGFR primary antibody (Biolegend, 345108), EGFR primary antibody (fisher scientific, MABF120MI), NT5E primary antibody (Biolegend, 344005), Alexa Fluor 488 Donkey Anti-Mouse secondary antibody (Jackson Labs, 715-545-151)

Drugs

Vemurafenib (Selleckchem, S1267), Dabrafenib (Cayman,16989-10), Trametinib (Cayman,16292-50), TGFB1 (R&D systems, 240-B-002), PI3K inhibitor (GDC-0941, Cayman, 11600-10) , TGFBRi (LY2109761, Sigma, SML2051-5MG)

Cell lines and tissue culture

We used the following cell lines: WM989 A6-G3, which are a twice single-cell bottlenecked clone of the melanoma line WM989 (provided by the Meenhard lab at the Wistar Institute), WM983B E9-D5, which are a twice single-cell bottlenecked clone of the melanoma line WM983B (provided by the Meenhard lab at the Wistar Institute), and HEK293FT cells which we used for lentiviral packaging. We authenticated the identity of all cell lines by STR profiling and confirmed that they are all negative for mycoplasma. STR profiling and mycoplasma testing was performed by the Penn Genomic Analysis Core. We cultured WM989 A6-G3 and WM983B E9-D5 in TU2% (78.4% MCDB 153, 19.6% Leibovitz's L-15, 2% FBS, 1.68mM CaCl, 50 Units/mL penicillin, and 50µg/mL streptomycin). We cultured HEK293FT in DMEM 5% (95% DMEM high glucose with GlutaMAX, 5% FBS, 50 Units/mL penicillin, and 50µg/mL streptomycin). We grew all cells at 37 °C and 5% CO_2., and pasaged all cells using 0.05% trypsin-EDTA.

Barcode Library

We used a high-complexity transcribed barcode library described in Emert et al. for our lineage barcodes (Emert et al. 2021). The plasmid uses LRG2.1T as a back bone, but we replaced the U6 promoter and sgRNA insert with the GFP sequence followed by a 100 nucleotide semi-random barcode, expressed by an EFS promoter. The barcode is semi-random as it is made up of WSN repeats (W = A or T, S = G or C, N = any) to maximize barcode complexity. A detailed protocol on the barcode production process can be found in Emert et al. which also links to this protocol: https://www.protocols.io/view/barcode-plasmid-library-cloning-4hggt3w.
The sequence of the plasmid can be found here:
https://benchling.com/s/seq-DAMUWPyU198hRSbpiecf?m=slm-GJ609ijArVWmkT8mk8zr

Lentiviral packaging

We grew HEK293FT to about 90% confluence in a 10cm dish containing 10mL of media (see the "cell lines and culture" section for details). To transfect cells with the lentiviral plasmid, we combined 500µl of OPTI-MEM with 80µl of 1mg/mL PEI in one tube. In a second tube, we combined 500µL of OPTI-MEM, 9 µg of the psPAX2 plasmid, 5.5µg of the VSVG plasmid, and 8µg of the barcode plasmid. Next, we combined the contents of these two tubes and let it incubate at room temperature for 15 minutes. After 15 minutes, we pipetted the solution dropwise into the plate of HEK293FT. We then incubated the cells at 37 °C for 7 hours. After 7 hours, we removed the media, washed the plate once with DPBS, and added 10ml of fresh media. We then incubated the cells at 37 °C for ~12 hours. After 12 hours, we checked that the cells were GFP positive with a fluorescence microscope to confirm that virus was produced. Next, we removed the media and put on a fresh 6mL of media. We then incubated the cells at 37 °C for ~12 hours. After 12 hours, we collected the media (the media contains the virus). We then repeated the process of adding 6mL of media and collecting virus every 12 hours for a total of ~72 hours. After the last collection, we filtered all the media containing the virus through a 0.2µm filter to ensure no HEK293FT cells were left with the virus media. Finally, we made 1,000µl aliquots of the media containing the virus and stored them at -80°C.

Lentiviral transduction

When barcoding cells we wanted to avoid multiple lineage barcodes per cell, for this reason we aimed to transduce ~20% of cells. To transduce the cells, we made a mixture of polybrene

(4µg/mL final concentration), virus (whatever amount was found to result in 20% transduction in a standard curve), and cells at 150,000 cells/mL. Next, we put 2 mL of this mixture into each well of a 6 well plate. We then spun the plate at 600 RCF for 25 minutes. Next, we incubated the cells with the virus at 37 °C for 8 hours. After the 8 hour incubation, we removed the media containing the virus, washed each well with DBPS, and added 2mL of fresh media to each well. The next day, we transferred each well to its own 10cm dish. We then gave the cells 2-3 days to start expressing the barcodes. We verified that barcodes were expressed by looking for GFP expression in the cells (cells express GFP along with the barcode).

Fluorescence-Activated Cell Sorting (FACS)

We dissociated cells into a single cell suspension using trypsin-EDTA and washed them once with 0.1% BSA. To stain for EGFR and NGFR, we would first stain with the EGFR antibody (see antibodies section) diluted 1:200 in 0.1% BSA for 1 hour on ice. We would then wash the cells twice with 0.1% BSA before staining with the anti-mouse A488 secondary antibody at a 1:500 dilution in 0.1% BSA for 30 minutes on ice. Next, to stain for NGFR, we washed the cells once with 0.1% BSA, and then resuspended in a 1:11 dilution of the NGFR antibody directly conjugated to APC in 0.5% BSA 2mM EDTA. We then incubated the cells on ice for 10 minutes. Finally we washed the cells once with 0.5% BSA 2mM EDTA, and then resuspended in 1% BSA with a low concentration of dapi and kept on ice until sorting.

To stain for NT5E, we resuspend the cells in a solution of NT5E antibody diluted 1:200 in .1% BSA and let the cells incubate for 30min. We then washed the cells twice with .1% BSA and then resuspended the cells in 1% BSA with a low concentration of dapi and kept on ice until sorting.

For flow sorting, we followed the staining protocols above and then sorted the cells on a Beckman Coulter Moflo Astrios with a 100µm nozzle. We used forward and side scatter to separate cells from debri and select singlets. We selected dapi negative cells to remove dead cells. To sort primed cells with EGFR and NGFR, we selected the top 0.2% of EGFR and NGFR expressing cells. To sort primed cells with NT5E, we selected the top 2% of NT5E expressing cells.

Single-cell RNA sequencing

We used 10X genomics 3' sequencing kit for all our scRNA-seq experiments. For the first scRNA-seq experiment, introduced in Figure 1 C, we sorted 1,000 barcoded WM989 cells per well in a 96 well plate (one well of mixed cells and one of EGFR/NGFR-high cells) and let them go through 4-5 doublings. We then trypsinized 1 mixed well and 1 primed well, and prepared as described in the chromium Single cell 3' Reagent kit V3 user guide. When loading cells on the microfluidics chip, we split both the primed and the mixed cells across 2 wells. After GEM generation, we continued to follow the chromium Single cell 3' Reagent kit V3 user guide to generate libraries.

For the second scRNA-seq experiment (data shown in Supplemental Figure 2 panels A and B), we sorted 1,000 WM989 primed cells (based on NT5E expression) and 1,000 mixed WM989 cells into one well of a 96 well plate, and 2,000 WM989B cells into another well. We then let the cells undergo 4 divisions. Once cells had divided 4 times, we trypsinized them and prepared the

samples all the way through library generation as described in the chromium Single cell 3' Reagent kit V3.1 (Dual Index) user guide.

For the third scRNA-seq experiment, shown in Figure 4, we sorted out 2,000 barcoded cells into a well of a 96 well plate. We let these cells divide for 7 to 8 divisions. We then split these cells across 4 plates, and gave them 1 day to adhere. We then started treatments for each plate (one plate untreated, one plate 5ng/mL TGFB1, one plate 4μM LY2109761 (TGFBRi), and one plate 2μM GDC-0941 (PI3K inhibitor)). We incubated the cells in their respective treatments for 5 days. We carried the above steps with two samples in parallel to have a replicate. After 5 days, we trypsinized the cells and processed them all the way through library generation as described in the chromium Single cell 3' Reagent kit V3.1 (Dual Index) user guide.

We sequence all our single-cell libraries using a NextSeq 500 with the 75 cycle high output kit (Illumina, 20024906). Samples sequenced with the Single cell 3' Reagent kit V3 used 8 reads for the index, 28 reads for read 1 and 49 reads for read 2. Samples sequenced with the Single cell 3' Reagent kit V3.1 (Dual Index) used 10 cycles for each index, 28 cycles for read 1 and 43 cycles for read 2.

Lineage barcode recovery from scRNA-seq

To recover the lineage barcodes we used some of the excess full length cDNA generated in the 10X protocol. Specifically, we used primers to selectively amplify the whole sequence between the 10X cell barcode and the end of the lineage barcode by PCR (Supp. Table 2). To perform the PCR, we used 100ng of full length cDNA per reaction, with 0.5μM of each primer, combined with PCR master mix (NEB, M0543S). We used 12 cycles to amplify the cDNA using the following protocol: an initial 30 second denature step at 98°C, then 98°C for 10 second, 65°C for 2 minutes repeated 12 times, followed by a 5 minute final extension step at 65°C. We then extract the amplified barcodes which are ~1.3kb using SPRI bead (Becman Coulter, B23317) size selection (we use a 0.6X bead concentration). To sequence the barcode library, we used a NextSeq 500 with a 150 cycle mid output kit (Illumina, 20024904). We did paired end sequencing and used 28 cycles on read 1 to read the 10X barcode and UMI, 8 cycles on each index, and 123 cycles on read 2 to get the lineage barcode sequence.

gDNA barcode recovery

To recover the barcodes from cells that we were not not performing scRNA-seq on, we extracted the barcodes from the gDNA. To do this, we trypsinized cells, pelleted them, and then extracted their gDNA using the QIAamp DNA Mini kit according to the manufacturer's protocol (Qiagen, 56304). To amplify the barcodes, we performed PCR using primers with homology to each side of the barcode. The primers also contain the illumina adapter sequence, and index sequences (see Supp. Table 2 for primer sequences). To perform the PCR amplification, we used 500ng isolated gDNA per reaction, 0.5μM of each primer, and PCR Master Mix (NEB, M0543S). We used 24 cycles to amplify the barcodes using the following protocol: an initial 30 second denature step at 98°C, then 98°C for 10 second, 65°C for 40 second repeated 24 times, followed by a 5 minute final extension step at 65°C. After amplification, we used SPRI beads (Becman Coulter, B23317) to select for the barcode which had an expected length of ~350bp. To isolate this

fragment size, we performed a two sided selection where we first select with a 0.6 X bead concentration and keep the supernatant (large gDNA fragments are on beads). We then select again using a 1.2X bead concentration keeping the material that binds to the beads this time (small fragments such as the primers are in the supernatant). To sequence the barcode library, we used a NextSeq 500 with a 150 cycle mid output kit (Illumina, 20024904). We performed single end sequencing and used 151 cycles on read 1 to read the lineage barcode, and then used 8 cycles on each index.

scRNA-seq analysis

We used the 10X Genomics Cellranger pipeline to generate fastq files (using the hg38 reference genome), generate the count matrix, and to aggregate replicate runs (without depth normalization). We also used the cellranger feature barcode pipeline to integrate out lineage barcodes with the scRNA-seq data (more information on how we did this in the "Combining single RNA sequencing and barcode data section").

Once we generated the aggregated count matrices with incorporated barcode information, we analyzed the data using Seurat V3. Using Seurat, we did basic filtering of the data based on the number of unique genes detected per cell, both removing poorly sequenced cells (low number of genes), and data points likely to be doublets (high number of genes). We also filtered based on the percent of mitochondrial reads to eliminate low quality or dying cells. If we saw batch effects between replicates, we used Seurat scRNA-seq integration pipeline to remove them. When there were no batch effects, we used SCtransform to normalize the data before running PCA, clustering, and dimensionality reduction with UMAP. When plotting gene expression information we did not use the SCtransform data, but rather separately log normalized the data. To generate scores per cell on how much they expressed a certain gene set we used the AUCell package (Aibar et al. 2017). AUCell accepts any gene set and then calculates the area under the curve for that gene set to determine if a gene set is enriched in a cell.

Combining scRNA-seq and barcode data

To recover lineage barcodes from the sequencing data, we used a custom python script (which can be found here: https://github.com/SydShafferLab/BarcodeAnalysis) combined with 10X Genomics CellRanger Feature Barcode pipeline. At a high level this pipeline works by identifying lineage barcode sequences in the FASTQ files and editing these FASTQ files so highly similar sequences are made to have identical sequences. All the unique barcodes are then put in a reference file which is used by Cell Ranger to identify the lineage barcode sequences in the FASTQ files and links the lineage barcode with the cell barcode.

More specifically, we identify lineage barcodes in the FASTQ files generated from lineage barcode sequencing by searching for a known sequence at the beginning of all lineage barcodes. Once we identified all potential barcode sequences, we used the STARCODE (https://github.com/gui11aume/starcode) package to identify barcodes that were very similar to each other and replace them all with the most frequently detected sequence within the set of similar barcodes. We then put these edited barcode sequences back into the fastq file, and generate a reference file containing all the edited barcode sequences. Next, we fed these edited

fastq files and the reference file into the cell ranger pipeline, and used the Feature Barcode analysis feature to link lineage barcodes with the cell barcodes. This provided us with the lineage and gene expression information for cells where a barcode was identified.

Our initial steps identifies real barcodes by combining similar barcodes, but when we looked at this output we found that we could more stringently call real lineages using additional filtering steps. The code used to accomplish this can be found in the "Assign a lineage to each cell" section of the 10X1_r1_r2_Analysis_unorm_sctrans.Rmd script. In brief, we first eliminated lineages that appear across multiple samples, as such lineages are not possible. We then also removed lineages that are bigger than expected given the amount of time cells were given to proliferate. Finally, for cells that appeared to have more than one lineage barcode, we tested whether there are multiple cells with this same combination of barcodes and considered those cells to be in the same lineage.

Validate primed marker

To validate if a cell surface protein is a marker for the primed state, we stained cells with a marker of interest, and then sorted the cells. Specifically, we sorted a mixed population of cells (sort all live cells) in one well, and a population of cells high in our marker of interest in another well (for sorting detail refer to the "Fluorescence-Activated Cell Sorting"section). We then gave the cells 24 hours to adhere to the plate before starting treatment with 1μM vemurafenib for 3 weeks. After 3 weeks in vemurafenib we counted the number of cells and drug resistant colonies in each well to determine if the marker increased the number of cells that survive targeted therapy.

PDX tumor generation

The tissue was collected from the Ashani Weeraratna lab and Arjun Raj lab (Torre et al. 2021). Briefly, these melanoma tumors were generated by subcutaneously injecting 1*10^6 WM989-A6-G3-Cas9-5a3 cells (These cells are WM989 A6-G3 that have been transduced with Cas9, and re-bottle necked) into 8-week old NOD/SCID mice. The mouse were fed AIN-76A chow, and the facilities were maintained between 21-23°C, a humidity of 30-35% and lights had a 12h on/off cycle with lights on from 6:00 to 18:00.The tumor was collected and fresh frozen when it measured ~1,500 mm^3.

Tissue RNA FISH

To probe for cells that are *NT5E*-high and *SOX10* low in the mouse PDX tumor, we used HCR V3 using probes for *NT5E* and *SOX10* (Acheampong et al. 2022; Choi et al. 2018). The probes for both genes were designed by molecular instruments (*NT5E* lot #: PRK825, *SOX10* lot #: PRK826), and we also bought the hairpin probes from molecular instruments. To perform HCR V3 in tissue we made slight modifications to the original protocol, therefore, we outline the whole protocol we used here. First, we used cryostat sectioning to generate 6μm sections of fresh frozen PDX tumor. We placed these sections on charged slides and fixed them with 4% formaldehyde for 10 minutes. We then washed the slides twice with 1X PBS and stored them in ethanol. To start HCR, we placed the slide in a slide staining tray and washed the slides twice with 5X SSC (sodium chloride sodium citrate). After removing the 5X SSC, we added 200μl of hybridisation buffer (30% formamide, 5x SSC, 9mM citric acid (pH 6.0), 0.1% Tween 20, 50ug/mL heparin, 1X Denhardt's

solution, 10% dextran sulfate) which was pre-heated to 37°C onto the tissue. We then incubated the slide for 10 minutes at 37°C. All incubation steps in this protocol were done with the slide staining tray closed and with water at the bottom to stop the sample from drying out. During this incubation period, we added 0.8 pmol of each probe pool (in this case *NT5E* and *SOX10*) to 200µl of probe hybridization buffer pre-heated at 37°C, and kept the solution at 37°C. After 10 minutes, we removed the hybridization buffer from the tissue and added the 300µl of hybridization buffer containing the probe pools. We placed a cover slip over the sample and incubated it for 12-16h at 37°C. After the incubation, we prepared the hair pins by putting 0.6 pmol each hair pin into its own tube (keep h1 and h2 separate), and put them into a thermocycler that heated up to 95°C for 90 seconds and then slowly cooled down to 25°C over 30 minutes. While the hair pins were being prepared in the thermo cycler, we added 300µl of wash buffer (30% formamide, 5x SSC, 9mM citric acid (pH 6.0), 0.1% Tween 20, 50ug/mL heparin) to the slide to help remove the cover slip. After we removed the coverslip, we performed multiple wash steps with decreasing amounts of wash buffer. We first added 300µl of 75% wash buffer, 25% 5xSSCT, removed that and added 300µl of 50% wash buffer, 50% 5xSSCT for 15 minutes, removed that and added 25% wash buffer, 75% 5xSSCT, and finally removed that and added 300µl of 100% 5X SSCT. For each step of the wash, we left the slides in each given solution for 15 minutes at 37°C. After the 100% 5X SSCT wash, we removed the 5X SSCT and added 200µl of room temperature amplification buffer (5X SSC, 0.1% Tween 20, 10% dextran sulfate) to the slide and incubated it at room temperature for 30 minutes. After 30 minutes, we removed the amplification buffer and added a solution of the prepared hairpins mixed with 100µl of amplification buffer to the slide, and added a cover slip. We incubated the slide in the staining tray at room temperature for 12-16 hours. After incubating with the hairpins, we removed the cover slip and washed the slide off using successive 5X SSC washes. We put 300µl of 5X SSC on the sample for 5 minutes, removed it, then added 5XSCC for 15 minutes, removed it, added 5X SSC for 15 minutes again, removed it, and finally added 5X SSC with Dapi for 5 minutes. After that, we mounted the slide using TrueVIEW (Vector labs, SP-8500-15), added a coverslip, and sealed it using nail polish.

Flow Cytometry

We dissociated cells from the plate using trypsin-EDTA into a single-cell suspension and washed once with 0.1%BSA. We then resuspended the cells in a 1:200 dilution of anti-NT5E antibody conjugated with APC and incubated them for 30min on ice. Next, we washed the cells once with 0.1% BSA, once with 1% BSA, and then resuspended them in 1% BSA for analysis by flow cytometry. We used an Accuri C6 for our flow cytometry, and quantified 10,000 events per sample. To analyze the data we used the R package flowCore (https://github.com/RGLab/flowCore). In our analysis we used forward and side scatter to identify cells, and used the FL4 channel (640nm excitation laser and 675/25 filter) to quantify cell surface levels of NT5E. To determine what percent of cells were primed, we set an intensity threshold where 2% of untreated cells would land above the threshold. We considered any cells above this threshold as primed.

Drug resistant colony experiments

We plated cells in 6 well plates with 10,000 cells per well. After plating, we gave cells 24 hours without treatment to adhere to the plate. After 24 hours, we initiated pre-treatments on wells that required it and would also change the media on no treatment controls. We incubated cells in their

respective treatment for 5 days. After 5 days we aspirated off the media and replaced it with media containing 250nM Dabrafenib and 2.5nM Trametinib. We maintained treatment with 250nM Dabrafenib and 2.5nM Trametinib for 4 weeks, changing the media every 3-4 days. After 4 weeks, we fixed the cells by aspirating off the media, washing the wells with DPBS, and treating them with 4% Formaldehyd for 10 minutes. We then aspirated off the formaldehyd, and washed twice with DBPS. After the washes we added 2mL of DPBS to each well, and stained the cells with DAPI. We then imaged the wells using a 10X objective on a confocal microscope.

Cell and colony counting

To determine how many cells are in wells after drug treatment, we use a pipeline that feeds tiled images using the custom pipeline DeepTile (https://github.com/arjunrajlaboratory/DeepTile/tree/071e3e9fb27f50ce024fd5ece25e3a4b0071f771) into DeepCell (https://github.com/vanvalenlab/deepcell-tf) to generate nuclear masks. To simplify the interface with DeepTile and DeepCell, as well as remove nuclei incorrectly called outside the well we use DeepCellHelper (https://github.com/SydShafferLab/DeepCellHelper). We then determine the number of cells per well by determining how many masks were generated per image. To identify colonies, we used a custom graphic user interface ColonySelector (https://github.com/SydShafferLab/ColonySelector) which uses the coordinates of each nucleus as input and allows users to circle colonies in each well, and saves a file containing which nucleus belongs to which colony. Using the output of the colony selecting software, we counted the number of colonies there were in each well.

Incucyte analysis

For time lapse experiments on the incucyte we used a clonal population of WM989 cells tagged with H2B-GFP for nuclear tracking. We took GFP images of the cells every 12 hours to track cell growth over time. We used the Incucyte software to generate nuclear masks, and exported csv tables containing the number of nuclei in each well at each time point. We then plotted this data in R.

Treatment growth effects

To determine if our treatments were leading to state specific changes in proliferation rates, we used WM989 cells tagged with H2B-GFP tagged cells and the incucyte to measure cell division over time. Specifically we sorted out primed cells and drug-susceptible cells into separate wells of a 96 well plate. After giving the cells 24 hours to adhere, we either added nothing, TGFB1, or PI3K inhibitor to both primed and drug-susceptible wells. We then imaged the cells using the incucyte as described in the “incucyte analysis” section above to determine the division rate of cells in each state and treatment combination.

Software and data availability

All data and code used for this paper can be found here:
https://drive.google.com/drive/folders/1-C78090Z43w5kGb1ZW8pXgysjha35jlU?usp=sharing

Author Contributions

G.H. and S.M.S. conceptualized the project and designed the study. G.H. performed all experiments and analysis with the following exceptions. R.A.R.H. helped with the barcode processing pipeline and provided helpful discussion. D.S. performed the ATAC sequencing experiment and its analysis, and helped with the barcode processing pipeline. M.S. and A.S performed the modeling analysis. B.E. provided technical guidance and troubleshooting for the barcoding library. A.T.W, M.E.F, and G.M.A generated the mouse PDX tissue. S.N. helped maintain cell lines and performed computational analyses. G.H. and S.M.S. wrote the paper.

Competing interests

S.M.S. receive royalties related to Stellaris RNA FISH probes. All other authors have no competing interests to declare.

Acknowledgements

We thank all members of the Shaffer lab for feedback on experiments and the manuscript, P. Gonzalez-Camara and T. Ridky for thoughtful discussions and ideas, W. Niu for help on developing image processing pipelines, L. Bugaj for feedback on the manuscript, and M. Herlyn for providing cell lines. S.M.S. acknowledges support from the NIH Director's Early Independence Award DP5OD028144 and the Wistar/Penn Skin Cancer SPORE (P50 CA174523).

Figure Captions

Figure 1: Cells primed for drug resistance have memory and exist in vivo A. Model for how the heritable primed state leads to resistance to targeted therapy. Cells primed for drug resistance (in green) proliferate and pass on their gene expression state through cell division, which demonstrates the concept of gene expression memory. These primed cells survive treatment with BRAFi and MEKi (resistant cells in red) while the drug-susceptible cells (in gray) die off. **B.** Schematic of lineages that maintain memory compared to those that lose memory. In lineages that maintain memory, the end state of all the cells is the same as the initial state. In lineages where memory is lost, the endstate comprises a mixture of cell states. **C.** Schematic of the experimental design that we used to capture the transcriptome and lineage of cells. We transduced melanoma cells (WM989) with the lentiviral lineage barcodes. We then sorted a sample of primed cells (based on their expression of EGFR and NGFR) and a mixed population. This sorting step provides us with the initial state of the cells. We then allowed the cells to undergo ~4 doublings to build out the lineages before doing scRNA-seq and barcode sequencing on the cells. **D.** UMAP plots of the per cell log10 normalized gene expression for the drug-susceptible associated marker *SOX10,* and the primed state associated markers *EGFR*, *AXL* and *NT5E*. **E.** Images of fixed NT5E-low and NT5E-high cells stained with DAPI after 3 weeks in targeted therapy. Each black dot is the nucleus of a drug-resistant cell, and drug-resistant colonies are circled in black. **F.** HCR RNA FISH of *NT5E* and *SOX10* in a mouse PDX drug naive tumor. Nuclei were stained with DAPI (blue) and RNA for each respective probe is in white. The green

arrow points to an example *SOX10*-low *NT5E*-high cell. The leftmost image is a large scan of the tissue section and the scale bar corresponds to 1mm, and the 2 images to the right are zoomed in on an example cell where the scale bar corresponds to 10μm.

Figure 2: An EMT-like state is activated early in the transition to the primed cell state A. Bar plots of the amount of cells and the size of each lineage type found in the data. Each bar represents a lineage, and the color of the bar indicates the state of the cell (green is primed and gray is drug-susceptible). In lineages that change state, the number of cells in each state is reflected by the colors in the bar. **B.** UMAP plots showing an example lineage from each type of lineage in the data. The cells from the example lineage are highlighted in blue. **C.** Schematic showing which cells from the drug-susceptible lineages (indicated in the blue rectangle) and which cells from lineages crossing from the drug-susceptible state to the primed state (indicated in the red rectangle) were used to identify genes important for transitioning. **D.** Volcano plot representing the differential expression analysis outlined in C. Red points represent genes up regulated in crossing lineages and blue points represent genes downregulated in crossing lineages**. E.** UMAP plot showing an integrated score of the activity for the gene signature defined by crossing lineage differential expression in D (high score means the cell highly expresses the gene set) **F.** UMAP plot labeling the top 2% drug-susceptible cells expressing the crossing lineage gene set. We classify these cells as “Intermediate cells”, and these cells are labeled in purple. **G.** Heatmap of the log10 normalized and scaled gene expression of cells in the drug-susceptible, intermediate, and primed state. **H.** Bar plot of the normalized enrichment score of the top 5 gene sets enriched in the crossing lineage gene set. **I.** UMAP plot showing which cells have high expression of the EMT gene set. Arrows point to drug-susceptible cells with high EMT scores. **J**. UMAP plot showing which cells have high expression of the TGF-β signaling gene set. Arrows point to drug-susceptible cells with high TGF-β signaling scores.

Figure 3: TGFB1 and PI3K inhibitor can modulate the number of primed cells in the population A. Schematic showing the expected number of primed cells over time if cells are untreated, treated with a drug or protein that induces the primed state, or treated with a drug or protein that disrupts the primed state. **B.** Example flow cytometry density plots from NT5E stained WM989 cells that have undergone 5 days of their respective treatments. **C.** Box plot of the log10 percent of WM989 cells in the primed state after 5 days in each respective treatment. All conditions have 3 biological replicates each with 6 technical replicates. P values were calculated using a wilcoxon test (ns: $p > 0.05$, *: $p <= 0.05$, **: $p <= 0.01$, ***: $p <= 0.001$, ****: $p <= 0.0001$). The dotted line represents the mean log10 percent of primed cells in the untreated condition. **D.** Box plot of the log10 percent of WM983B cells in the primed state after 5 days of each respective treatment. All conditions have 2 biological replicates each with 3 technical replicates. P values were calculated using a wilcoxon test (**: $p <= 0.01$). The dotted line marks the mean log10 percent of primed cells in the untreated condition.

Figure 4: Treatment with TGFB1 induces the primed state and treatment with PI3K inhibitor induces the drug-susceptible state. A. Schematic of the experimental design used to determine if treatments cause state switching. We transduced melanoma cells (WM989) with lineage

barcodes, allowed them to divide ~7 times, and then split them across 4 plates, expecting each plate to get cells with the same barcode. Cells with the same barcodes serve as effectively as "copies" of the same cells, therefore, by treating each barcode with all the different treatments we can determine which treatment causes state switching. We used scRNA-seq with barcode sequencing to capture both the lineage and transcriptome of the cells at the end point. **B.** UMAP plots of the log10 normalized gene expression of the drug-susceptible state associated marker *SOX10* and primed state associated markers *EGFR*, *AXL* and *NT5E*. **C.** UMAP plots highlighting cells in each condition relative to all the other sequenced cells (in gray). Cells highlighted in blue were untreated, cells highlighted in orange were treated with TGFB1, cells highlighted in teal were treated with TGFBRi, and cells highlighted in pink were treated with the PI3Ki. **D.** UMAP plot with the primed cells labeled in green and the drug-susceptible cells in gray. **E**. Bar plot quantifying the log10 percent of primed cells in each condition based on the defined drug-susceptible and primed cell populations in panel D and the location of each treatment condition shown in panel C. P values were calculated using chi squared tests (ns: $p > 0.05$, ****: $p <= 0.0001$) **F.** Heatmaps of the log10 normalized and scaled gene expression of drug-susceptible and primed cells in each treatment condition.

Figure 5: Changes in the number of primed cells occur through state switching A. Schematic of lineage based analysis to test for state switching into the primed state. Box plots show the fraction of cells in each lineage that are in the primed state across all the conditions. The lineages shown on this plot are exclusively those that were completely drug-susceptible in the untreated sample. P values were calculated using a wilcoxon test (ns: $p > 0.05$, *: $p <= 0.05$, **: $p <= 0.01$, ***: $p <= 0.001$, ****: $p <= 0.0001$) **B.** Schematic of lineage based analysis to test for state switching into the drug-susceptible state. Box plots show the fraction of cells in each lineage that are in the primed state across all the conditions. The lineages shown on this plot are exclusively those that were completely primed in the untreated sample. P values were calculated using a wilcoxon test (ns: $p > 0.05$, *: $p <= 0.05$, **: $p <= 0.01$) **C.** Schematic of the variables tested in our model to identify different ways changes in the number of primed cells can occur. The parameters tested include: drug-susceptible state death and proliferation rate, primed state death and proliferation rate, the transition rate from the drug-susceptible state to the primed state (k on), and the transition rate from the primed state to the drug-susceptible state (k off) **D.** Scatter plot showing the fraction of primed cells per lineage in the untreated control (x axis) relative to the TGFB1 treated cells (y axis). The underlying heatmap shows the results of 1 million simulations of what the data looks like if TGFB1 were to increase k_{on} such that the more primed cells there are in your lineage the larger your kon value.

Figure 6: Reduction in the number of primed cells results in less drug resistance A. Schematic of the experimental design testing whether modulators of state switching affect the number of drug-resistant colonies (represented by red cells). In this experiment we treated cells with the state switching modulating treatments for 5 days, removed the treatment, and then treated with BRAFi and MEKi for 4 weeks. After 4 weeks we quantified drug resistance by counting the number of drug-resistant cells and the number of drug-resistant colonies. **B.** Example images of fixed cells stained with dapi for each condition after 5 days of pretreatment and 4 weeks of selection with BRAFi and MEKi. Drug-resistant colonies are circled in black where readily

identifiable. Scale bar in the first image represents 5mm and applies for all images in the panel. **C.** Box plot quantifying the number of drug-resistant cells from scans like those shown in panel C across 3 biological replicates, each with 6 technical replicates. P values were calculated using a wilcoxon test (ns: $p > 0.05$, *: $p <= 0.05$, **: $p <= 0.01$, ***: $p <= 0.001$, ****: $p <= 0.0001$) **D.** Box plot quantifying the number of drug colonies (for each condition where distinct colonies were present) from scans like those shown in panel C across 3 biological replicates, each with 6 technical replicates. P values were calculated using a wilcoxon test (ns: $p > 0.05$, *: $p <= 0.05$, **: $p <= 0.01$, ***: $p <= 0.001$, ****: $p <= 0.0001$)

References

Aibar, Sara, Carmen Bravo González-Blas, Thomas Moerman, Vân Anh Huynh-Thu, Hana Imrichova, Gert Hulselmans, Florian Rambow, et al. 2017. "SCENIC: Single-Cell Regulatory Network Inference and Clustering." *Nature Methods* 14 (11): 1083–86.

Angel, Andrew, Jie Song, Caroline Dean, and Martin Howard. 2011. "A Polycomb-Based Switch Underlying Quantitative Epigenetic Memory." *Nature* 476 (7358): 105–8.

Bakin, Tomlinson, and Bhowmick. n.d. "Phosphatidylinositol 3-Kinase Function Is Required for Transforming Growth Factor β-Mediated Epithelial to Mesenchymal Transition and Cell Migration." *Bollettino Della Societa Italiana Di Biologia Sperimentale*. https://www.jbc.org/article/S0021-9258(20)88586-X/abstract.

Barkley, Dalia, Reuben Moncada, Maayan Pour, Deborah A. Liberman, Ian Dryg, Gregor Werba, Wei Wang, et al. 2021. "Recurrence of Cancer Cell States across Diverse Tumors and Their Interactions with the Microenvironment." *bioRxiv*. https://doi.org/10.1101/2021.12.20.473565.

Bedard, Philippe L., Josep Tabernero, Filip Janku, Zev A. Wainberg, Luis Paz-Ares, Johan Vansteenkiste, Eric Van Cutsem, et al. 2015. "A Phase Ib Dose-Escalation Study of the Oral Pan-PI3K Inhibitor Buparlisib (BKM120) in Combination with the Oral MEK1/2 Inhibitor Trametinib (GSK1120212) in Patients with Selected Advanced Solid Tumors." *Clinical Cancer Research*. https://doi.org/10.1158/1078-0432.ccr-14-1814.

Bhang, Hyo-Eun C., David A. Ruddy, Viveksagar Krishnamurthy Radhakrishna, Justina X. Caushi, Rui Zhao, Matthew M. Hims, Angad P. Singh, et al. 2015. "Studying Clonal Dynamics in Response to Cancer Therapy Using High-Complexity Barcoding." *Nature Medicine* 21 (5): 440–48.

Biddy, Brent A., Wenjun Kong, Kenji Kamimoto, Chuner Guo, Sarah E. Waye, Tao Sun, and Samantha A. Morris. 2018. "Single-Cell Mapping of Lineage and Identity in Direct Reprogramming." *Nature* 564 (7735): 219–24.

Chaligne, Ronan, Federico Gaiti, Dana Silverbush, Joshua S. Schiffman, Hannah R. Weisman, Lloyd Kluegel, Simon Gritsch, et al. 2021. "Epigenetic Encoding, Heritability and Plasticity of Glioma Transcriptional Cell States." *Nature Genetics* 53 (10): 1469–79.

Das Thakur, Meghna, Fernando Salangsang, Allison S. Landman, William R. Sellers, Nancy K. Pryer, Mitchell P. Levesque, Reinhard Dummer, Martin McMahon, and Darrin D. Stuart. 2013. "Modelling Vemurafenib Resistance in Melanoma Reveals a Strategy to Forestall Drug Resistance." *Nature* 494 (7436): 251–55.

Derynck, Rik, and Ying E. Zhang. 2003. "Smad-Dependent and Smad-Independent Pathways in TGF-Beta Family Signalling." *Nature* 425 (6958): 577–84.

Emert, Benjamin L., Christopher J. Cote, Eduardo A. Torre, Ian P. Dardani, Connie L. Jiang, Naveen Jain, Sydney M. Shaffer, and Arjun Raj. 2021. "Variability within Rare Cell States Enables Multiple Paths toward Drug Resistance." *Nature Biotechnology*, February. https://doi.org/10.1038/s41587-021-00837-3.

Filipp, Fabian V., Chen Li, and Alexander D. Boiko. 2019. "CD271 Is a Molecular Switch with Divergent Roles in Melanoma and Melanocyte Development." *Scientific Reports* 9 (1): 7696.

Gupta, Piyush B., Christine M. Fillmore, Guozhi Jiang, Sagi D. Shapira, Kai Tao, Charlotte Kuperwasser, and Eric S. Lander. 2011. "Stochastic State Transitions Give Rise to Phenotypic Equilibrium in Populations of Cancer Cells." *Cell* 146 (4): 633–44.

Hoek, Keith S., Ossia M. Eichhoff, Natalie C. Schlegel, Udo Döbbeling, Nikita Kobert, Leo Schaerer, Silvio Hemmi, and Reinhard Dummer. 2008. "In Vivo Switching of Human Melanoma Cells between Proliferative and Invasive States." *Cancer Research* 68 (3): 650–56.

Irvine, Mal, Ashleigh Stewart, Bernadette Pedersen, Suzanah Boyd, Richard Kefford, and Helen Rizos. 2018. “Oncogenic PI3K/AKT Promotes the Step-Wise Evolution of Combination BRAF/MEK Inhibitor Resistance in Melanoma.” *Oncogenesis* 7 (9): 72.

Jardim, Denis L., Débora De Melo Gagliato, Mina Nikanjam, Donald A. Barkauskas, and Razelle Kurzrock. 2020. “Efficacy and Safety of Anticancer Drug Combinations: A Meta-Analysis of Randomized Trials with a Focus on Immunotherapeutics and Gene-Targeted Compounds.” *Oncoimmunology* 9 (1): 1710052.

Jerby-Arnon, Livnat, Parin Shah, Michael S. Cuoco, Christopher Rodman, Mei-Ju Su, Johannes C. Melms, Rachel Leeson, et al. 2018. “A Cancer Cell Program Promotes T Cell Exclusion and Resistance to Checkpoint Blockade.” *Cell* 175 (4): 984–97.e24.

Kaur, Amanpreet, Karun Kiani, Dylan Fingerman, Margaret C. Dunagin, Jingxin Li, Ian Dardani, Eric M. Sanford, et al. 2022. “Metastatic Potential in Clonal Melanoma Cells Is Driven by a Rare, Early-Invading Subpopulation.” *bioRxiv*. https://doi.org/10.1101/2022.04.17.488591.

Kavran, Andrew J., Scott A. Stuart, Kristyn R. Hayashi, Joel M. Basken, Barbara J. Brandhuber, and Natalie G. Ahn. 2022. “Intermittent Treatment of BRAFV600E Melanoma Cells Delays Resistance by Adaptive Resensitization to Drug Rechallenge.” *Proceedings of the National Academy of Sciences of the United States of America* 119 (12): e2113535119.

La Manno, Gioele, Ruslan Soldatov, Amit Zeisel, Emelie Braun, Hannah Hochgerner, Viktor Petukhov, Katja Lidschreiber, et al. 2018. “RNA Velocity of Single Cells.” *Nature* 560 (7719): 494–98.

Lee, Jenny H., Elena Shklovskaya, Su Yin Lim, Matteo S. Carlino, Alexander M. Menzies, Ashleigh Stewart, Bernadette Pedersen, et al. 2020. “Transcriptional Downregulation of MHC Class I and Melanoma de- Differentiation in Resistance to PD-1 Inhibition.” *Nature Communications* 11 (1): 1897.

Mendoza, Michelle C., E. Emrah Er, and John Blenis. 2011. “The Ras-ERK and PI3K-mTOR Pathways: Cross-Talk and Compensation.” *Trends in Biochemical Sciences* 36 (6): 320–28.

Meng, Lingbin, Baoqiong Liu, Rui Ji, Xin Jiang, Xuebo Yan, and Ying Xin. 2019. “Targeting the BDNF/TrkB Pathway for the Treatment of Tumors.” *Oncology Letters* 17 (2): 2031–39.

Neal, James T., Xingnan Li, Junjie Zhu, Valeria Giangarra, Caitlin L. Grzeskowiak, Jihang Ju, Iris H. Liu, et al. 2018. “Organoid Modeling of the Tumor Immune Microenvironment.” *Cell* 175 (7): 1972–88.e16.

Neftel, Cyril, Julie Laffy, Mariella G. Filbin, Toshiro Hara, Marni E. Shore, Gilbert J. Rahme, Alyssa R. Richman, et al. 2019. “An Integrative Model of Cellular States, Plasticity, and Genetics for Glioblastoma.” *Cell* 178 (4): 835–49.e21.

Nymark Aasen, Synnøve, Himalaya Parajuli, Tuyen Hoang, Zichao Feng, Krister Stokke, Jiwei Wang, Kislay Roy, Rolf Bjerkvig, Stian Knappskog, and Frits Thorsen. 2019. “Effective Treatment of Metastatic Melanoma by Combining MAPK and PI3K Signaling Pathway Inhibitors.” *International Journal of Molecular Sciences* 20 (17): 4235.

Oren, Yaara, Michael Tsabar, Michael S. Cuoco, Liat Amir-Zilberstein, Heidie F. Cabanos, Jan-Christian Hütter, Bomiao Hu, et al. 2021. “Cycling Cancer Persister Cells Arise from Lineages with Distinct Programs.” *Nature* 596 (7873): 576–82.

Park, Sook Ryun, Myrtle Davis, James H. Doroshow, and Shivaani Kummar. 2013. “Safety and Feasibility of Targeted Agent Combinations in Solid Tumours.” *Nature Reviews. Clinical Oncology* 10 (3): 154–68.

Pedri, Dennis, Panagiotis Karras, Ewout Landeloos, Jean-Christophe Marine, and Florian Rambow. 2022. “Epithelial-to-Mesenchymal-like Transition Events in Melanoma.” *The FEBS Journal* 289 (5): 1352–68.

Perotti, V., P. Baldassari, A. Molla, C. Vegetti, I. Bersani, A. Maurichi, M. Santinami, A. Anichini, and R. Mortarini. 2016. “NFATc2 Is an Intrinsic Regulator of Melanoma Dedifferentiation.” *Oncogene* 35 (22): 2862–72.

Qiu, Qi, Peng Hu, Xiaojie Qiu, Kiya W. Govek, Pablo G. Cámara, and Hao Wu. 2020.

"Massively Parallel and Time-Resolved RNA Sequencing in Single Cells with scNT-Seq." *Nature Methods* 17 (10): 991–1001.
Radke, Josefine, Florian Roßner, and Torben Redmer. 2017. "CD271 Determines Migratory Properties of Melanoma Cells." *Scientific Reports* 7 (1): 9834.
Rafsanjani Nejad, Pouria, Pradip Shahi Thakuri, Sunil Singh, Astha Lamichhane, Jacob Heiss, and Hossein Tavana. 2021. "Toxicity of Combinations of Kinase Pathway Inhibitors to Normal Human Cells in a Three-Dimensional Culture." *SLAS Technology* 26 (3): 255–64.
Raghavan, Srivatsan, Peter S. Winter, Andrew W. Navia, Hannah L. Williams, Alan DenAdel, Kristen E. Lowder, Jennyfer Galvez-Reyes, et al. 2021. "Microenvironment Drives Cell State, Plasticity, and Drug Response in Pancreatic Cancer." *Cell* 184 (25): 6119–37.e26.
Rambow, Florian, Aljosja Rogiers, Oskar Marin-Bejar, Sara Aibar, Julia Femel, Michael Dewaele, Panagiotis Karras, et al. 2018. "Toward Minimal Residual Disease-Directed Therapy in Melanoma." *Cell* 174 (4): 843–55.e19.
Ramsdale, Rachel, Robert N. Jorissen, Frederic Z. Li, Sheren Al-Obaidi, Teresa Ward, Karen E. Sheppard, Patricia E. Bukczynska, et al. 2015. "The Transcription Cofactor c-JUN Mediates Phenotype Switching and BRAF Inhibitor Resistance in Melanoma." *Science Signaling* 8 (390): ra82.
Shaffer, Sydney M., Margaret C. Dunagin, Stefan R. Torborg, Eduardo A. Torre, Benjamin Emert, Clemens Krepler, Marilda Beqiri, et al. 2017. "Rare Cell Variability and Drug-Induced Reprogramming as a Mode of Cancer Drug Resistance." *Nature* 546 (7658): 431–35.
Shaffer, Sydney M., Benjamin L. Emert, Raúl A. Reyes Hueros, Christopher Cote, Guillaume Harmange, Dylan L. Schaff, Ann E. Sizemore, et al. 2020. "Memory Sequencing Reveals Heritable Single-Cell Gene Expression Programs Associated with Distinct Cellular Behaviors." *Cell*, July. https://doi.org/10.1016/j.cell.2020.07.003.
Shapiro, Geoffrey I., Patricia LoRusso, Eunice Kwak, Susan Pandya, Charles M. Rudin, Carla Kurkjian, James M. Cleary, et al. 2020. "Phase Ib Study of the MEK Inhibitor Cobimetinib (GDC-0973) in Combination with the PI3K Inhibitor Pictilisib (GDC-0941) in Patients with Advanced Solid Tumors." *Investigational New Drugs*. https://doi.org/10.1007/s10637-019-00776-6.
Sharma, Sreenath V., Diana Y. Lee, Bihua Li, Margaret P. Quinlan, Fumiyuki Takahashi, Shyamala Maheswaran, Ultan McDermott, et al. 2010. "A Chromatin-Mediated Reversible Drug-Tolerant State in Cancer Cell Subpopulations." *Cell* 141 (1): 69–80.
Shimizu, Toshio, Anthony W. Tolcher, Kyriakos P. Papadopoulos, Muralidhar Beeram, Drew W. Rasco, Lon S. Smith, Shelly Gunn, et al. 2012. "The Clinical Effect of the Dual-Targeting Strategy Involving PI3K/AKT/mTOR and RAS/MEK/ERK Pathways in Patients with Advanced Cancer." *Clinical Cancer Research: An Official Journal of the American Association for Cancer Research* 18 (8): 2316–25.
Simiczyjew, Aleksandra, Katarzyna Pietraszek-Gremplewicz, Ewelina Dratkiewicz, Marta Podgórska, Rafał Matkowski, Marcin Ziętek, and Dorota Nowak. 2019. "Combination of Selected MET and EGFR Inhibitors Decreases Melanoma Cells' Invasive Abilities." *Frontiers in Pharmacology*. https://doi.org/10.3389/fphar.2019.01116.
Stommel, Jayne M., Alec C. Kimmelman, Haoqiang Ying, Roustem Nabioullin, Aditya H. Ponugoti, Ruprecht Wiedemeyer, Alexander H. Stegh, et al. 2007. "Coactivation of Receptor Tyrosine Kinases Affects the Response of Tumor Cells to Targeted Therapies." *Science* 318 (5848): 287–90.
Sun, Chong, Liqin Wang, Sidong Huang, Guus J. J. E. Heynen, Anirudh Prahallad, Caroline Robert, John Haanen, et al. 2014. "Reversible and Adaptive Resistance to BRAF(V600E) Inhibition in Melanoma." *Nature* 508 (7494): 118–22.
Tang, Yaqi, Simon Durand, Stéphane Dalle, and Julie Caramel. 2020. "EMT-Inducing Transcription Factors, Drivers of Melanoma Phenotype Switching, and Resistance to Treatment." *Cancers* 12 (8). https://doi.org/10.3390/cancers12082154.

Tirosh, Itay, Benjamin Izar, Sanjay M. Prakadan, Marc H. Wadsworth 2nd, Daniel Treacy, John J. Trombetta, Asaf Rotem, et al. 2016. “Dissecting the Multicellular Ecosystem of Metastatic Melanoma by Single-Cell RNA-Seq.” *Science* 352 (6282): 189–96.

Torre, Eduardo A., Eri Arai, Sareh Bayatpour, Connie L. Jiang, Lauren E. Beck, Benjamin L. Emert, Sydney M. Shaffer, et al. 2021. “Genetic Screening for Single-Cell Variability Modulators Driving Therapy Resistance.” *Nature Genetics* 53 (1): 76–85.

Verfaillie, Annelien, Hana Imrichova, Zeynep Kalender Atak, Michael Dewaele, Florian Rambow, Gert Hulselmans, Valerie Christiaens, et al. 2015. “Decoding the Regulatory Landscape of Melanoma Reveals TEADS as Regulators of the Invasive Cell State.” *Nature Communications* 6 (April): 6683.

Villanueva, Jessie, Adina Vultur, John T. Lee, Rajasekharan Somasundaram, Mizuho Fukunaga-Kalabis, Angela K. Cipolla, Bradley Wubbenhorst, et al. 2010. “Acquired Resistance to BRAF Inhibitors Mediated by a RAF Kinase Switch in Melanoma Can Be Overcome by Cotargeting MEK and IGF-1R/PI3K.” *Cancer Cell* 18 (6): 683–95.

Wegiel, Barbara, Anders Bjartell, Zoran Culig, and Jenny Liao Persson. 2007. “Interleukin-6 Activates PI3K/Akt Pathway and Regulates Cyclin A1 to Promote Prostate Cancer Cell Survival.” *International Journal of Cancer*. https://doi.org/10.1002/ijc.23261.

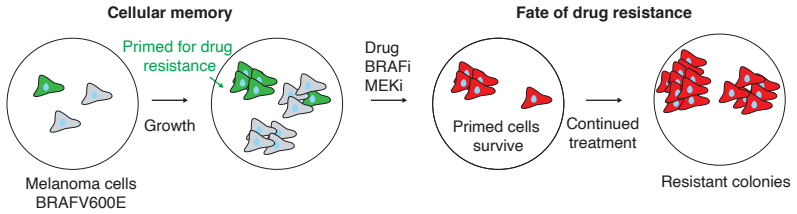
Weinreb, Caleb, Alejo Rodriguez-Fraticelli, Fernando D. Camargo, and Allon M. Klein. 2020. “Lineage Tracing on Transcriptional Landscapes Links State to Fate during Differentiation.” *Science* 367 (6479). https://doi.org/10.1126/science.aaw3381.

Xu, Jian, Samy Lamouille, and Rik Derynck. 2009. “TGF-Beta-Induced Epithelial to Mesenchymal Transition.” *Cell Research* 19 (2): 156–72.

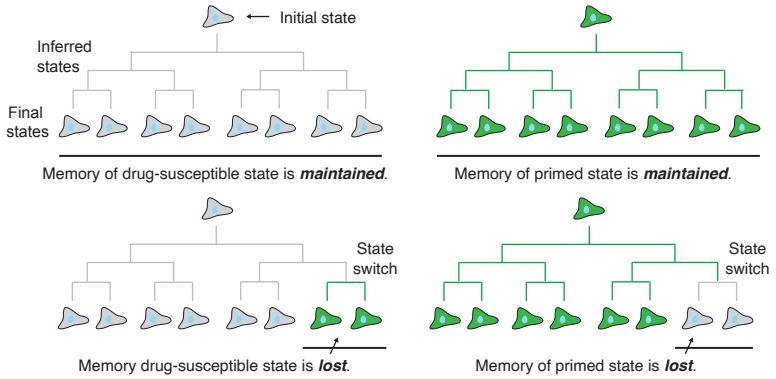
Zhang, Ying E. 2009. “Non-Smad Pathways in TGF-Beta Signaling.” *Cell Research* 19 (1): 128–39.

Figure 1

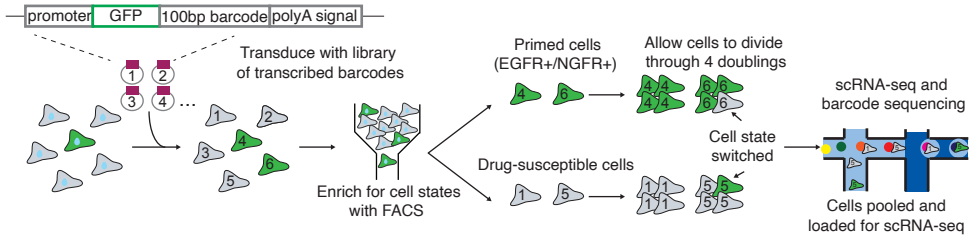
A



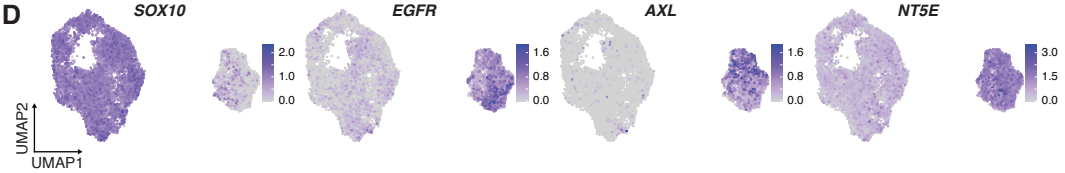
B **Cellular memory is revealed by the end state of cell lineages.**



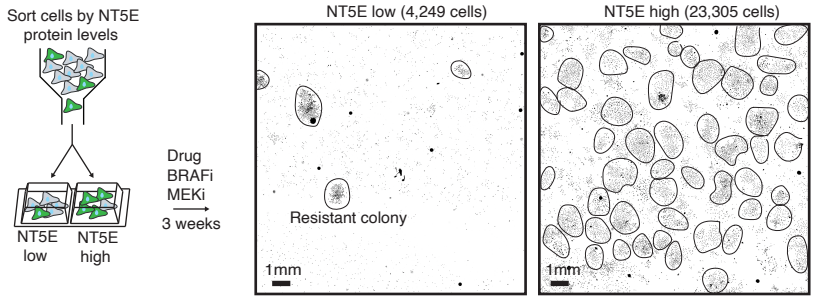
C **scMemorySeq combines high-throughput cell barcoding and scRNA-seq to measure cellular memory.**



D



E **NT5E is a marker of cells that are primed for drug resistance.**



F RNA FISH HCRv3.0

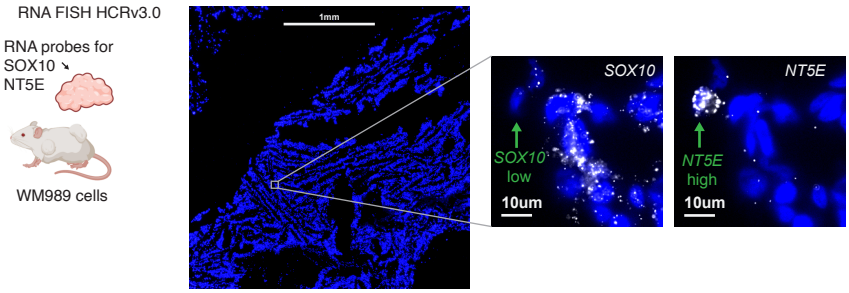


Figure 2

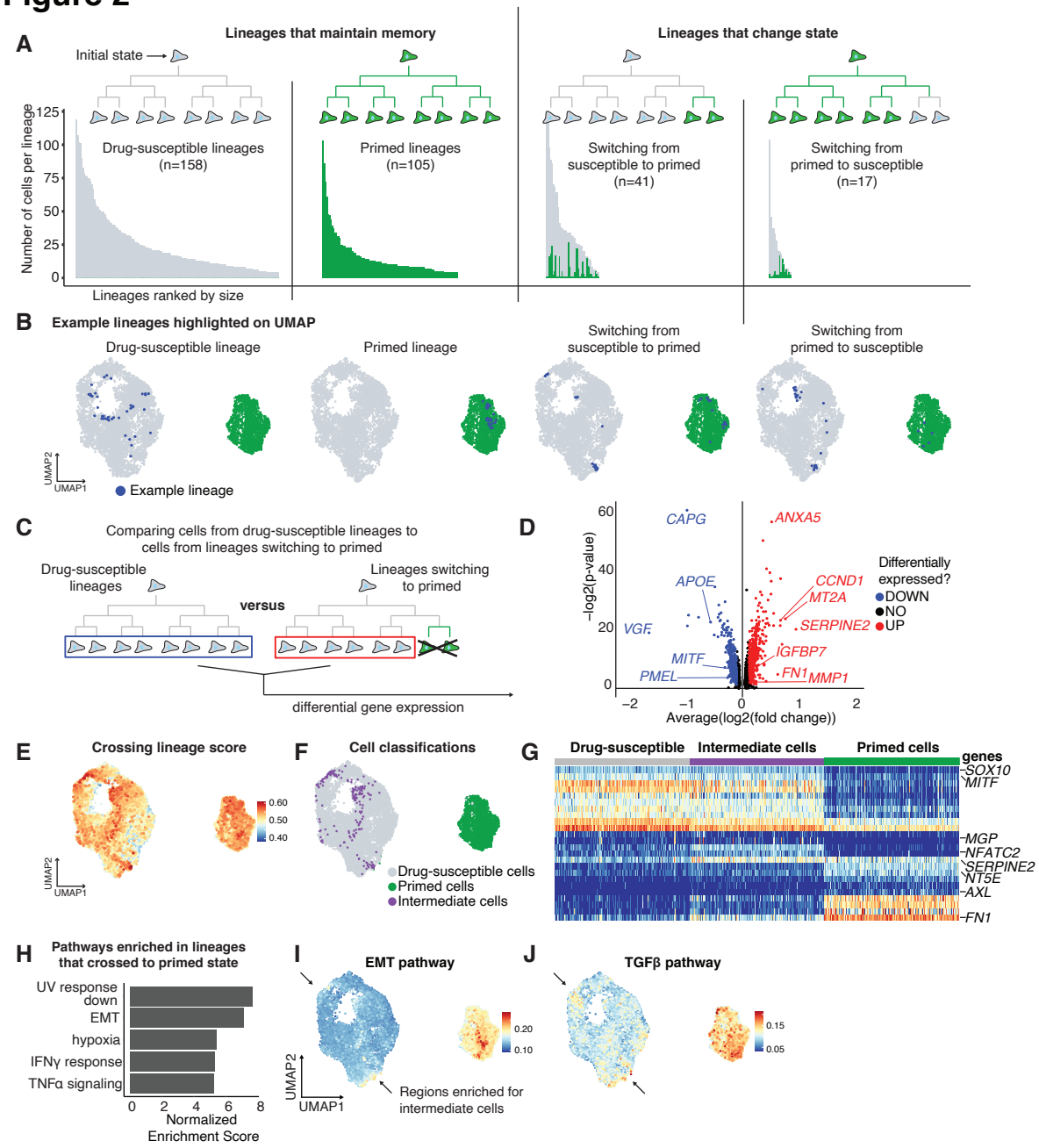

A
Lineages that maintain memory
Lineages that change state
Initial state
Number of cells per lineage
Drug-susceptible lineages (n=158)
Primed lineages (n=105)
Switching from susceptible to primed (n=41)
Switching from primed to susceptible (n=17)
Lineages ranked by size
B
Example lineages highlighted on UMAP
Drug-susceptible lineage
Primed lineage
Switching from susceptible to primed
Switching from primed to susceptible
Example lineage
C
Comparing cells from drug-susceptible lineages to cells from lineages switching to primed
Drug-susceptible lineages
Lineages switching to primed
versus
differential gene expression
D
CAPG
ANXA5
APOE
CCND1
MT2A
SERPINE2
VGF
IGFBP7
MITF
FN1
PMEL
MMP1
Differentially expressed?
DOWN
NO
UP
−log2(p-value)
Average(log2(fold change))
E
Crossing lineage score
F
Cell classifications
Drug-susceptible cells
Primed cells
Intermediate cells
G
Drug-susceptible
Intermediate cells
Primed cells
genes
SOX10
MITF
MGP
NFATC2
SERPINE2
NT5E
AXL
FN1
H
Pathways enriched in lineages that crossed to primed state
UV response down
EMT
hypoxia
IFNγ response
TNFα signaling
Normalized Enrichment Score
I
EMT pathway
Regions enriched for intermediate cells
J
TGFβ pathway

Figure 3

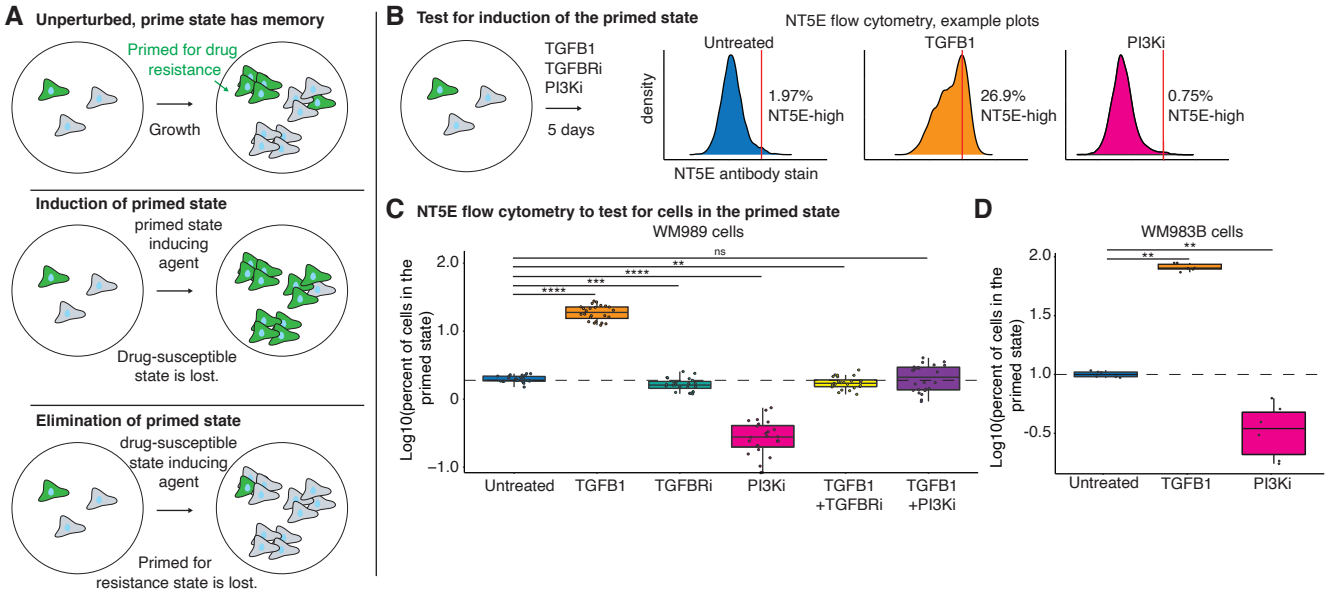

A
Unperturbed, prime state has memory
Primed for drug resistance
Growth
Induction of primed state
primed state inducing agent
Drug-susceptible state is lost.
Elimination of primed state
drug-susceptible state inducing agent
Primed for resistance state is lost.
B
Test for induction of the primed state
TGFB1
TGFBRi
PI3Ki
5 days
NT5E flow cytometry, example plots
Untreated
TGFB1
PI3Ki
density
1.97%
NT5E-high
26.9%
NT5E-high
0.75%
NT5E-high
NT5E antibody stain
C
NT5E flow cytometry to test for cells in the primed state
WM989 cells
Log10(percent of cells in the primed state)
2.0
1.0
0
−1.0
ns
**
****
***
****
Untreated
TGFB1
TGFBRi
PI3Ki
TGFB1 +TGFBRi
TGFB1 +PI3Ki
D
WM983B cells
Log10(percent of cells in the primed state)
2.0
1.5
1.0
-0.5
**
**
Untreated
TGFB1
PI3Ki

Figure 4

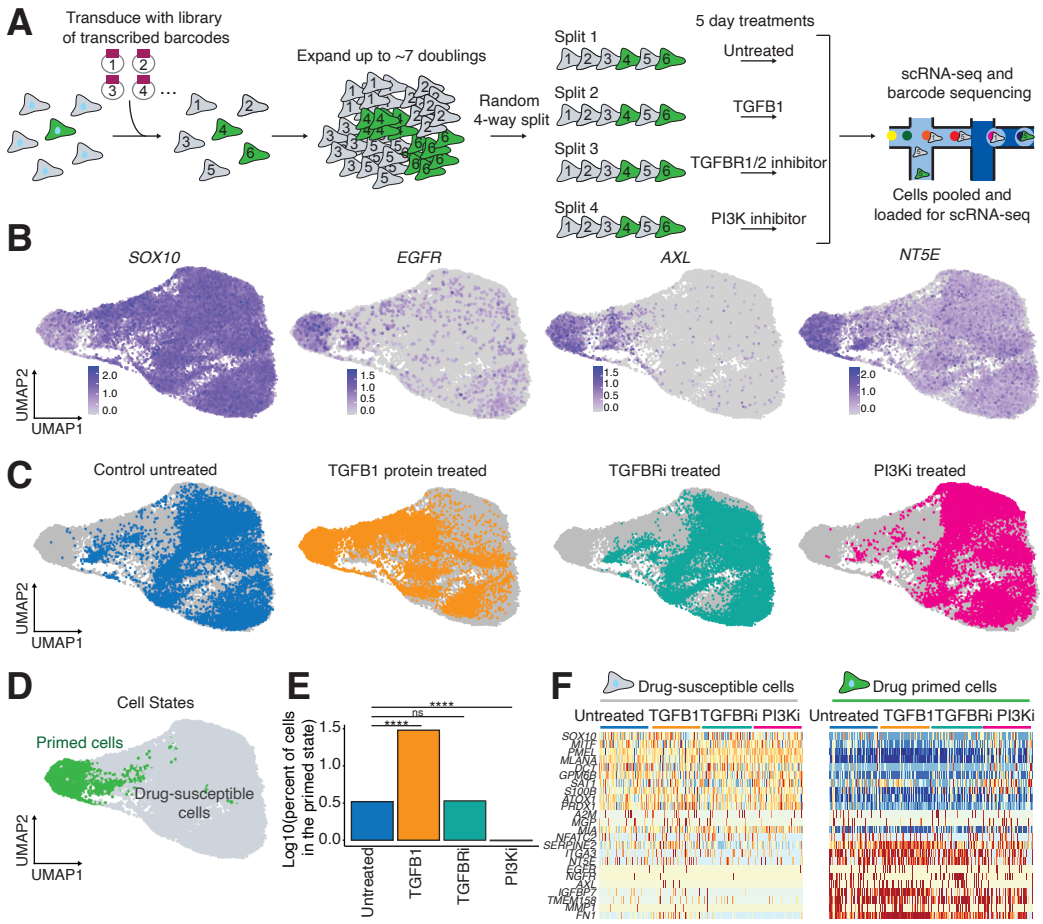
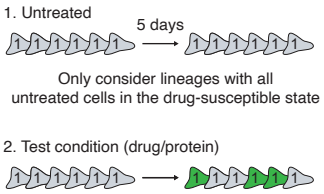
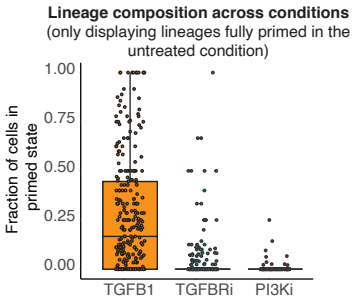

A
Transduce with library of transcribed barcodes
Expand up to ~7 doublings
Random 4-way split
Split 1
Split 2
Split 3
Split 4
5 day treatments
Untreated
TGFB1
TGFBR1/2 inhibitor
PI3K inhibitor
scRNA-seq and barcode sequencing
Cells pooled and loaded for scRNA-seq
B
SOX10
EGFR
AXL
NT5E
UMAP2
UMAP1
C
Control untreated
TGFB1 protein treated
TGFBRi treated
PI3Ki treated
D
Cell States
Primed cells
Drug-susceptible cells
E
Log10(percent of cells in the primed state)
Untreated
TGFB1
TGFBRi
PI3Ki
ns
****
F
Drug-susceptible cells
Drug primed cells
Untreated TGFB1 TGFBRi PI3Ki
SOX10
MITF
PMEL
MLANA
DCT
GPM6B
SAT1
S100B
BLOX1
PRDX1
A2M
MGP
MIA
NFATC2
SERPINE2
ITGA3
NT5E
EGFR
NGFR
AXL
IGFBP7
TMEM158
MMP1
FN1

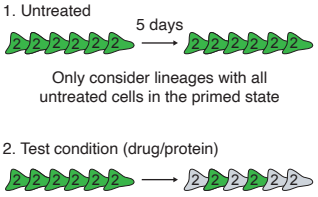
Figure 5

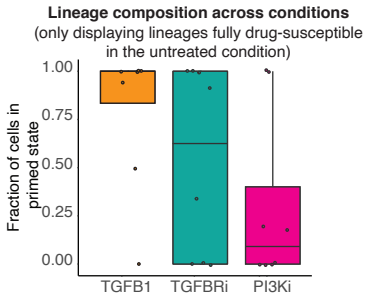
A Testing for **switching into the primed state**





B Testing for **switching out of the primed state**





C **Parameters governing fraction of cells in each state**

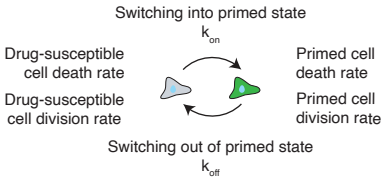
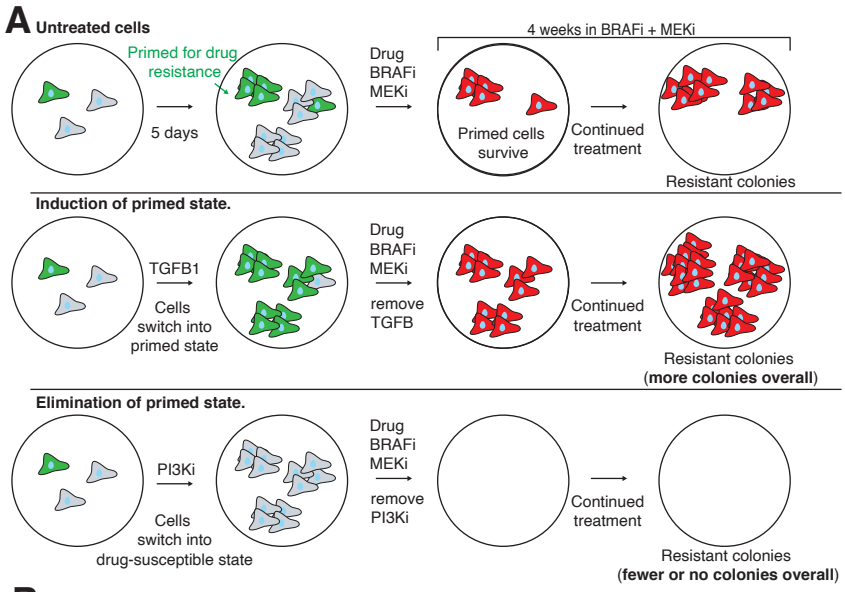
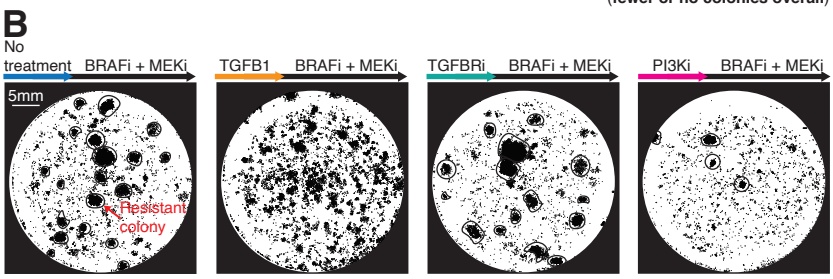
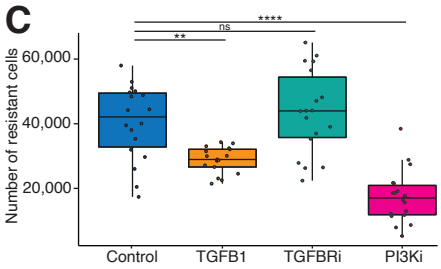
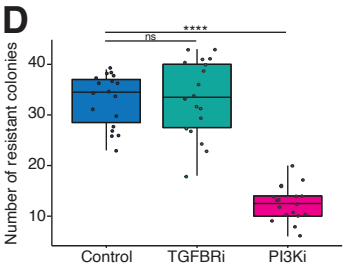


Figure 6


A
Untreated cells
Primed for drug resistance
5 days
Drug
BRAFi
MEKi
4 weeks in BRAFi + MEKi
Primed cells survive
Continued treatment
Resistant colonies
Induction of primed state.
TGFB1
Cells switch into primed state
Drug
BRAFi
MEKi
remove TGFB
Continued treatment
Resistant colonies
(more colonies overall)
Elimination of primed state.
PI3Ki
Cells switch into drug-susceptible state
Drug
BRAFi
MEKi
remove PI3Ki
Continued treatment
Resistant colonies
(fewer or no colonies overall)


B
No treatment
BRAFi + MEKi
TGFB1
BRAFi + MEKi
TGFBRi
BRAFi + MEKi
PI3Ki
BRAFi + MEKi
5mm
Resistant colony


C
Number of resistant cells
60,000
40,000
20,000
ns
****
**
Control
TGFB1
TGFBRi
PI3Ki


D
Number of resistant colonies
40
30
20
10
****
ns
Control
TGFBRi
PI3Ki