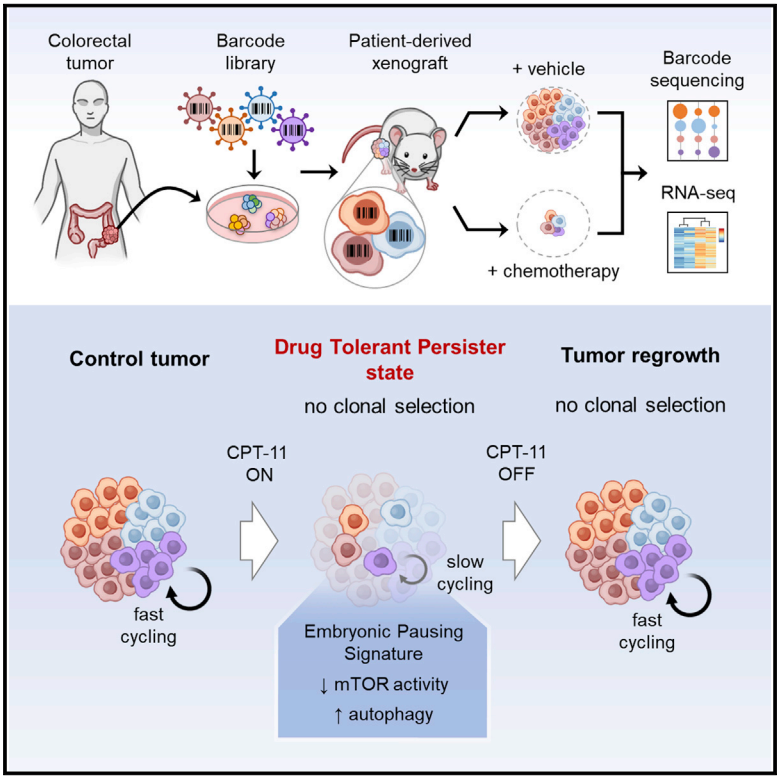
Cell

Colorectal Cancer Cells Enter a Diapause-like DTP State to Survive Chemotherapy

Graphical Abstract



Highlights

- CRC cells possess an equipotent capacity to enter a drug-tolerant persister state
- Tumors that recur following a DTP state maintain clonal complexity
- DTP-state tumors are similar to diapause, an embryonic survival program
- Similar to diapause, DTPs are dependent on autophagy for survival

Authors

Sumaiyah K. Rehman, Jennifer Haynes, Evelyne Collignon, ..., Sidhartha Goyal, Jason Moffat, Catherine A. O'Brien

Correspondence

mrsantos@lunenfeld.ca (M.R.-S.), goyal@physics.utoronto.ca (S.G.), j.moffat@utoronto.ca (J.M.), cobrien@uhnresearch.ca (C.A.O.)

In Brief

Any cancer cell has the ability to enter a drug-tolerant persister state in response to chemotherapy regimens by acquiring a reversible functional state akin to diapause.

Rehman et al., 2021, Cell *184*, 226–242
January 7, 2021 © 2020 Elsevier Inc.
https://doi.org/10.1016/j.cell.2020.11.018

CellPress

CellPress

Cell

Article

Colorectal Cancer Cells Enter a Diapause-like DTP State to Survive Chemotherapy

Sumaiyah K. Rehman,[1,15] Jennifer Haynes,[1,15,16] Evelyne Collignon,[2,15] Kevin R. Brown,[3] Yadong Wang,[1] Allison M.L. Nixon,[4] Jeffrey P. Bruce,[1] Jeffrey A. Wintersinger,[3,5,6,10] Arvind Singh Mer,[1,7] Edwyn B.L. Lo,[1] Cherry Leung,[1] Evelyne Lima-Fernandes,[1] Nicholas M. Pedley,[1] Fraser Soares,[1] Sophie McGibbon,[8] Housheng Hansen He,[1,7] Aaron Pollet,[2] Trevor J. Pugh,[5,7,9] Benjamin Haibe-Kains,[1,6,7] Quaid Morris,[3,4,5,6,10,11] Miguel Ramalho-Santos,[2,4,*] Sidhartha Goyal,[8,12,*] Jason Moffat,[3,4,12,*] and Catherine A. O'Brien[1,5,13,14,17,*]

[1]Princess Margaret Cancer Center, University Health Network, Toronto, ON M5G 1L7, Canada
[2]Lunenfeld-Tanenbaum Research Institute, Mount Sinai Hospital, Toronto, ON M5T 3L9, Canada
[3]Donnelly Centre, University of Toronto, Toronto, ON M5S 3E1, Canada
[4]Department of Molecular Genetics, University of Toronto, Toronto, ON M5S 1A8, Canada
[5]Ontario Institute for Cancer Research, Toronto, ON M5G 0A3, Canada
[6]Department of Computer Science, University of Toronto, Toronto, ON M5T 3A1, Canada
[7]Department of Medical Biophysics, University of Toronto, Toronto, ON M5G 1L7, Canada
[8]Department of Physics, University of Toronto, Toronto, ON M5S 1A7, Canada
[9]Clinical Genomics Program, Princess Margaret Cancer Centre, University Health Network, Toronto, ON M5G 2C1, Canada
[10]Vector Institute, Toronto, ON M5G 1M1, Canada
[11]Computational and Systems Biology Program, Memorial Sloan Kettering Cancer Center, New York, NY 10065, USA
[12]Institute of Biomaterials and Biomedical Engineering, University of Toronto, Toronto, ON M5S 3E1, Canada
[13]Department of Laboratory Medicine and Pathobiology, University of Toronto, Toronto, ON M5S 1A8, Canada
[14]Department of Surgery, University Health Network, Toronto, ON M5G 1L7, Canada
[15]These authors contributed equally
[16]Present address: Department of Clinical and Translational Sciences, Joan C. Edwards School of Medicine, Marshall University, Huntington, WV, USA
[17]Lead Contact
*Correspondence: mrsantos@lunenfeld.ca (M.R.-S.), goyal@physics.utoronto.ca (S.G.), j.moffat@utoronto.ca (J.M.), cobrien@uhnresearch.ca (C.A.O.)
https://doi.org/10.1016/j.cell.2020.11.018

SUMMARY

Cancer cells enter a reversible drug-tolerant persister (DTP) state to evade death from chemotherapy and targeted agents. It is increasingly appreciated that DTPs are important drivers of therapy failure and tumor relapse. We combined cellular barcoding and mathematical modeling in patient-derived colorectal cancer models to identify and characterize DTPs in response to chemotherapy. Barcode analysis revealed no loss of clonal complexity of tumors that entered the DTP state and recurred following treatment cessation. Our data fit a mathematical model where all cancer cells, and not a small subpopulation, possess an equipotent capacity to become DTPs. Mechanistically, we determined that DTPs display remarkable transcriptional and functional similarities to diapause, a reversible state of suspended embryonic development triggered by unfavorable environmental conditions. Our study provides insight into how cancer cells use a developmentally conserved mechanism to drive the DTP state, pointing to novel therapeutic opportunities to target DTPs.

INTRODUCTION

There is increasing evidence that non-genetic processes drive drug tolerance, presenting a major hurdle to successful cancer therapy (Recasens and Munoz, 2019). Drug-tolerant persister (DTP) cells are emerging as key players in the field of non-genetic heterogeneity of tumors and have been identified across a wide range of tumors in response to chemotherapy and targeted agents (Guler et al., 2017; Hangauer et al., 2017; Liau et al., 2017; Sharma et al., 2010). Therefore, DTPs represent a potential therapeutic opportunity prior to development of classic irreversible genetic mutation-driven drug resistance. The DTP state is characterized by cancer cells that are quiescent or slow-cycling and typically represent a small subpopulation of the parental tumor (0.3%–5%) (Guler et al., 2017; Liau et al., 2017; Sharma et al., 2010). The concept of persisters is derived from the microbial literature, where it is well established that treatment of infections with antibiotics can reduce bacterial burden but, in some instances, fails to eliminate refractory bacteria (Kaldalu and Tenson, 2019; Van den Bergh et al., 2017). Bacterial persisters are

characterized as phenotypic variants that can transiently tolerate extraordinary levels of antibiotics but remain genetically drug sensitive (Balaban et al., 2004, 2013; Brauner et al., 2016; Kint et al., 2012). In the context of cancer, DTPs represent a reversible state where, upon removal of treatment, often referred to as a drug holiday, they resume growth and proliferation but remain sensitive to chemotherapy (Recasens and Munoz, 2019). Similar to bacterial persisters, the ability of DTPs to escape the toxic effects of chemotherapy cannot be explained on the basis of genetic mutations (Recasens and Munoz, 2019). Pioneering studies by Sharma et al. (2010) identified DTPs in cancer cell lines and demonstrated that this drug-tolerant subpopulation could be ablated selectively *in vitro* by treatment with an insulin-like growth factor 1 receptor (IGF-1R) inhibitor or chromatin-modifying agents. Other studies have also demonstrated a role for repression of stress-induced long interspersed repeat element 1 (LINE-1) expression in driving the DTP state, as well as DTP dependency on the lipid hydroperoxidase GPX4 (Guler et al., 2017; Hangauer et al., 2017). To date, the majority of DTP studies have focused on *in vitro*-generated DTPs from commercial cell lines, demonstrating that they exist as a small subset of the total tumor cell population (Egan et al., 2015; Guler et al., 2017). More recent work in triple-negative breast cancer, using patient-derived xenografts (PDXs), also revealed that only a subpopulation of cancer cells was capable of repopulating the tumor following a DTP state (Echeverria et al., 2019). Interestingly, tumor regrowth was not associated with any genomic selection; therefore, it was suggested that phenotypic properties may have endowed a subpopulation of cancer cells with the capacity to preferentially re-populate the tumor. However, much remains to be understood about the subset of cancer cells that can generate DTPs and the key mechanisms driving the DTP state *in vivo*.

Given that DTPs represent a non-genetic reversible state cancer cells enter to survive the harsh environment created by chemotherapy,we questioned whether tumors are hijacking an evolutionarily conserved survival strategy. There are limited numbers of conserved survival strategies organisms employ to survive hostile environments, and most are conserved across the animal kingdom, including torpor, estivation, and diapause. Torpor and estivation are typically activated for an animal to survive environments of extreme cold and heat, respectively. Diapause is defined as a reversible state of suspended embryonic development triggered by unfavorable environmental conditions, including nutrient deprivation (Deng et al., 2018; Fenelon and Renfree, 2018). Diapause is a physiological reproductive strategy utilized across the animal kingdom, including in some mammals, to survive stressful environments. There is no cellular selection in diapause; rather, every cell in the embryo reversibly enters the state, and when the environmental insult resolves, the embryo resumes normal development. We investigated whether cancer cells, and more specifically DTPs, enter an embryonic diapause-like state to survive the stressful environment created by chemotherapy.

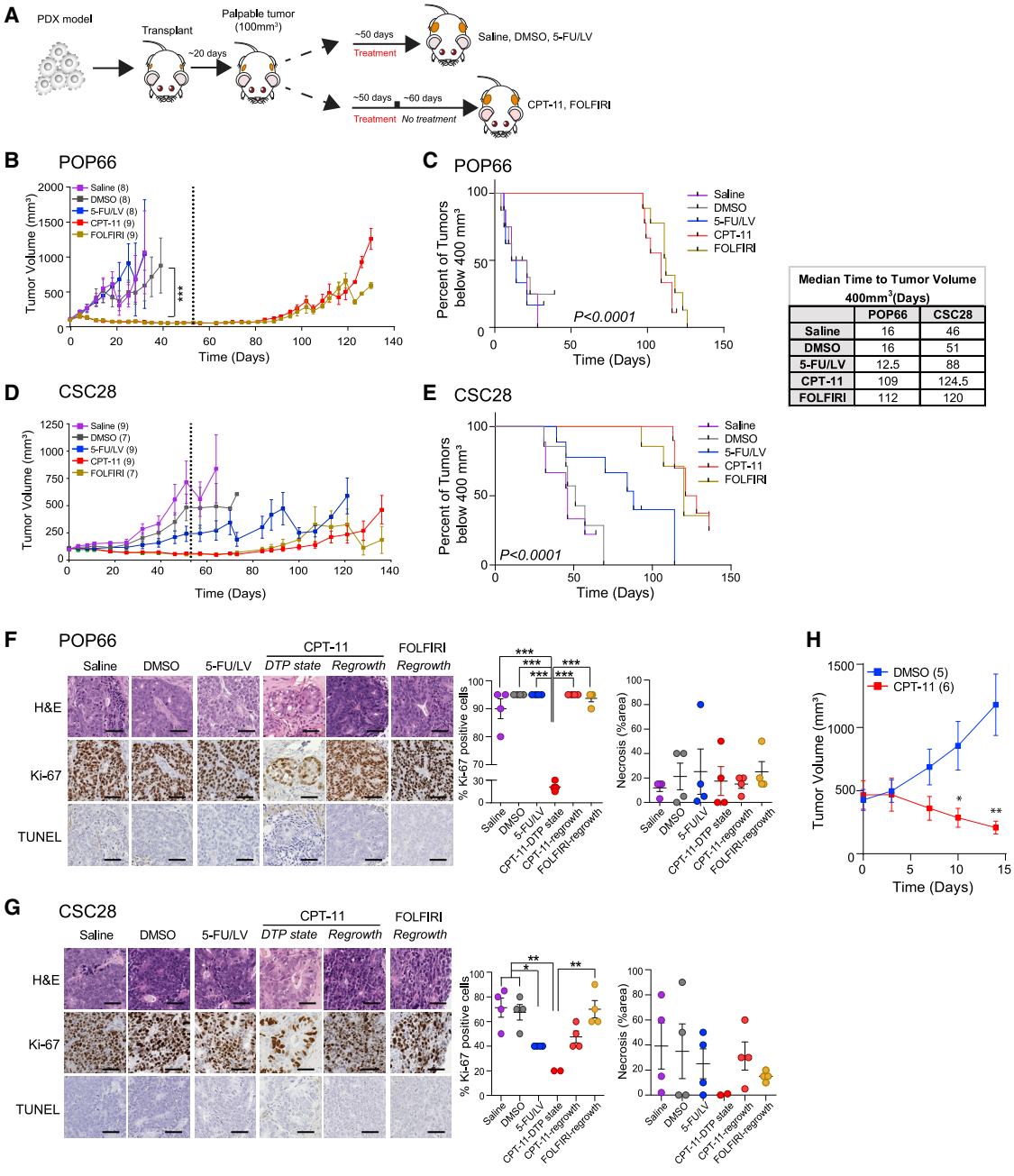
Here we used colorectal cancer (CRC) patient-derived models, high-complexity lentiviral barcoding, next-generation sequencing (NGS), and mathematical modeling to identify drivers of the DTP state. We found that xenograft tumors that emerged following a chemotherapy-induced DTP state did not exhibit any significant decrease in genetic or barcode complexity. Furthermore, mathematical modeling supported the "equipotent" model, in which all cancer cells in a given tumor have an equal capacity to enter the DTP state to survive therapy. Indeed, transcriptional profiling revealed that tumor cells adopt a state similar to the diapause state in embryos to survive the stress of chemotherapy and, similar to the embryo, can exit this state and reestablish the tumor upon stress removal (Deng et al., 2018; Fenelon and Renfree, 2018).

RESULTS

Response of Patient-Derived CRC Models to *In Vivo* Treatment with Standard-of-Care Chemotherapy

To study DTPs *in vivo*, we established a robust model that generates DTP-state tumors. First, we determined the effect of treatment with standard-of-care chemotherapy on xenograft tumor growth using two patient-derived CRC samples in immunodeficient mice (non-obese diabetic [NOD]/severe combined immunodeficiency [SCID]) (Table S1; Figures 1A and S1A). When tumors reached ~100 mm^3, mice were treated with standard-of-care chemotherapy regimens for CRC, using maximum tolerated doses (Figures S1B and S1C) (Gustavsson et al., 2015). For the patient-derived CRC model POP66, tumors from mice treated with 5-fluorouracil and leucovorin (5-FU/LV), oxaliplatin (OXA), or FOLFOX (5-FU/LV and OXA) demonstrated a small but not statistically significant growth delay compared with the vehicle control (saline; Figures S1D and S1E). Similarly, for CRC model CSC28, there was no significant tumor growth delay of OXA- and FOLFOX-treated tumors compared with saline, although 5-FU/LV-treated tumors had a small but significant growth delay (Figures S1F and S1G). There was no difference in Ki-67 expression in all treatment groups for POP66; however, CSC28 demonstrated a significant decrease in Ki-67 expression in the 5-FU/LV- and FOLFOX-treated groups (Figures S1H and S1I). No significant differences in TUNEL (terminal deoxynucleotidyl transferase dUTP nick end labeling) positivity or necrosis were observed across all tumor groups for both models (Figures S1H and S1J). The response to 5-FU/LV, OXA, or FOLFOX treatment, which ranged from minimal to no delay in tumor growth, was likely due to dose limitations of the NOD/SCID mice used in these studies.

In contrast, tumors (POP66 and CSC28) from mice treated with irinotecan (CPT-11) or FOLFIRI (5-FU/LV and CPT-11) for 8 weeks demonstrated negligible growth compared with the vehicle controls (saline for 5-FU/LV and 5% DMSO for CPT-11; Figures 1B–1E). Tumors remained stable for the 8 weeks during which mice were on treatment. However, for both CRC samples, tumor regrowth initiated when treatment with CPT-11 or FOLFIRI was discontinued (between 31–44 days for POP66 and 34–57 days for CSC28; Figures 1B and 1D). The median time for tumors to reach 400 mm^3 was delayed significantly in the CPT-11 and FOLFIRI groups compared with vehicle controls or 5-FU/LV treatment alone (Figures 1C and 1E). There were no significant differences in Ki-67 expression, TUNEL positivity, or percent necrosis between vehicle controls, CPT-11, and FOLFIRI regrowth tumors (Figures 1F and 1G). However, tumors that were



Median Time to Tumor Volume 400mm³(Days)

	POP66	CSC28
Saline	16	46
DMSO	16	51
5-FU/LV	12.5	88
CPT-11	109	124.5
FOLFIRI	112	120

Figure 1. Response of CRC PDXs to Treatment with Chemotherapy

(A) *In vivo* treatment schematic for two CRC patient-derived xenograft (PDX) models (POP66 and CSC28).

(B and C) Tumor growth curves (B) and Kaplan-Meier survival curves (C) of POP66 PDXs treated with saline, DMSO, 5-FU/LV, CPT-11, or FOLFIRI.

(D and E) Tumor growth curves (D) and Kaplan-Meier survival curves (E) of CSC28 PDXs treated with saline, DMSO, 5-FU/LV, CPT-11, or FOLFIRI. Numbers in parentheses indicate biological replicates in that group. A dotted line indicates when treatment was stopped. Each point on the growth curve is the mean tumor volume ± SEM, two-way ANOVA. The median time to reach a tumor volume of 400 mm^3 is listed for the POP66 and CSC28 PDXs; log rank test.

(F and G) Immunohistochemistry (H&E, Ki-67, and TUNEL stain, 20×) analysis of tumors for POP66 (F) from (B) and CSC28 (G) from (D) at the endpoint. CPT-11 DTP-state tumors were harvested on day 28 of treatment. Scale bar, 50 μm. Percent Ki-67 positive cells and percent area necrosis are plotted; one-way ANOVA.

(H) Reinjection of CPT-11-treated tumors that regrew (B, CPT-11 regrowth tumors) and re-treatment with DMSO or CPT-11. Numbers in parentheses indicate biological replicates in that group; t test.

*p < 0.05, **p < 0.01, ***p < 0.001. See also Table S1 and Figures S1 and S2.

harvested while on treatment with CPT-11 demonstrated significantly less Ki-67 expression with no significant alterations in percent necrosis or apoptosis, indicating decreased proliferation (Figures 1F and 1G). To determine whether CRC cells treated with CPT-11 or FOLFIRI remained sensitive following treatment cessation and tumor regrowth, we harvested viable tumor cells from the regrowth tumors and injected them into new mice, which were subsequently treated with CPT-11 (Figure 1H). All tumors remained sensitive to CPT-11 upon re-exposure and again demonstrated no growth while on treatment. Thus, the tumor response to CPT-11 and FOLFIRI was consistent with the cancer cells entering a reversible DTP state. In addition, limiting dilution assays on control (DMSO) versus CPT-11-DTP-state tumors showed no enrichment of stem cells in the DTP state (data not shown), establishing that it is not stem cells that preferentially survive this state.

Long-Term CPT-11 Treatment Gives Rise to Irreversibly Resistant Tumors

To generate tumors that were irreversibly resistant to CPT-11 treatment, we continued long-term *in vivo* exposure in a subset of mice from both CRC samples (Figure S2A). The CPT-11 dose was well tolerated by mice over the entire course of treatment (Figures S2B and S2C). For POP66, continuous long-term treatment *in vivo* resulted in three CPT-11-resistant tumors (of 10 total) that grew on treatment (~214 days; Figures S2D and S2E). Interestingly, despite being on treatment for ~15 months (469 days), resistant tumors never developed for CSC28 (Figures S2F and S2G). POP66 CPT-11-resistant tumors showed robust Ki-67 staining, indicating active tumor growth while on treatment (Figure S2H). Furthermore, POP66 and CSC28 long-term-treated tumors that did not grow on treatment also showed strong Ki-67 staining, suggesting that the tumor cells were cycling (Figure S2H). To determine whether the POP66 CPT-11 resistant tumors remained insensitive to therapy, we harvested viable tumor cells from the resistant tumors and injected them into new mice that were then treated with CPT-11 (Figure S2I). All three resistant tumors regrew while on treatment with CPT-11, indicating establishment of tumors that were irreversibly resistant, unlike tumors that went through a DTP state and regrew when CPT-11 treatment was stopped.

Genetic Heterogeneity Is Maintained from Parent Tumor to Treated Xenografts

We conducted whole-exome sequencing (WES) to determine the genetic clonal and sub-clonal structure of our CRC models. Analysis of the single-nucleotide variant (SNV) allele fractions (VAFs) in diploid regions of the POP66 parent tumor (P0), first-generation xenograft (G0), and CPT-11 regrowth xenografts indicated the presence of subclones in the parent tumor and xenografts (Figures S3A and S3B). Furthermore, by utilizing the mutant allele tumor heterogeneity (MATH) score to measure the genetic intra-tumoral heterogeneity (ITH), we determined that there was no significant change in genetic ITH observed between POP66 P0, G0, and CPT-11 regrowth xenografts (Figure S3B). Similarly for CSC28, genetic ITH was maintained from G0 through to the treated xenografts (saline, DMSO, CPT-11 regrowth; Figure S3C). MATH scores revealed that the POP66 CPT-11-resistant xenografts exhibited a significant decrease in genetic ITH compared with P0, G0, and CPT-11 regrowth tumors (Figure S3B), suggesting that they underwent genetic selection.

To further assess genetic ITH, we used the extensive mutational information from the POP66 and CSC28 WES data and subjected it to Pairtree computational analysis (Data S1; Wintersinger et al., 2020; Dobson et al., 2020). Using SNVs from diploid regions across all of the POP66 or CSC28 set of tumors (Data S1), we resolved unique subclones (POP66 = 26, CSC28 = 11) to construct clone trees describing the evolution of each cancer as a whole. We resolved a complex tree structure that demonstrated pervasive genetic ITH, with the proportions of each clone showing that every tumor sample (P0, G0, treated xenografts) is composed of multiple genetically distinct subpopulations (Figures S3D–S3G; Figure 1 in Data S1).

To compare the degree of genetic ITH between individual tumors, we analyzed each tumor independently using Pairtree (Figure 2 in Data S1;). For each tumor under different treatment conditions, we resolved the genetic subclones present, the relationships between these clones, the proportions of each clone in the tumor, and degree of genetic ITH indicated by the mutational Shannon diversity index (MSDI; Figure 2 in Data S1; Figure S3H) (Wolf et al., 2019). MSDI analysis supported our finding of pervasive ITH across POP66 and CSC28 P0, G0, vehicle, and CPT-11 regrowth xenografts and less genetic ITH in POP66 CPT-11-resistant xenografts (Figure S3H).

Our data establish that CPT-11-treated tumors that go through the DTP state do not demonstrate loss of genetic heterogeneity, unlike CPT-11-resistant tumors. Moreover, the absence of a decrease in genetic heterogeneity in the CPT-11 regrowth tumors, together with their retained sensitivity to CPT-11, is consistent with a non-genetic mechanism of drug tolerance.

Barcode Complexity Is Maintained in CRC Xenografts following a CPT-11-Induced DTP State

To probe the clonal dynamics of the DTP state *in vivo*, we adapted a barcoding strategy to track individual cells within tumors (Bhang et al., 2015). Briefly, prior to injection, we infected cells of the two CRC patient-derived models described above with a high-complexity lentiviral barcode library (~2.6 million variable DNA sequences or "barcodes") at a low multiplicity of infection (MOI $\leq$ 0.1) to limit the number of multiple viral integrants. Following selection and *in vitro* expansion, cells were injected subcutaneously into the flanks of NOD/SCID mice (Figure 2A). A decrease in overall barcode complexity was observed following xenograft generation, where an engraftment efficiency of 5%–15% was determined based on barcode abundance in the expanded cell population used for the injections (T0) and the number of barcoded cells injected (Table S2). Using spiked-in control barcodes, we confirmed that NGS can robustly quantify barcodes representing as few as 50 cells. The barcodes that were present at higher abundance at the time of injection into mice (T0) were generally not the most abundant barcodes in xenografts (Figures 2B, 2C, S4A, and S4B), indicating no bias on the basis of initial barcode abundance.

To elucidate the effect of the DTP state on clonal complexity, barcode composition was compared between control and

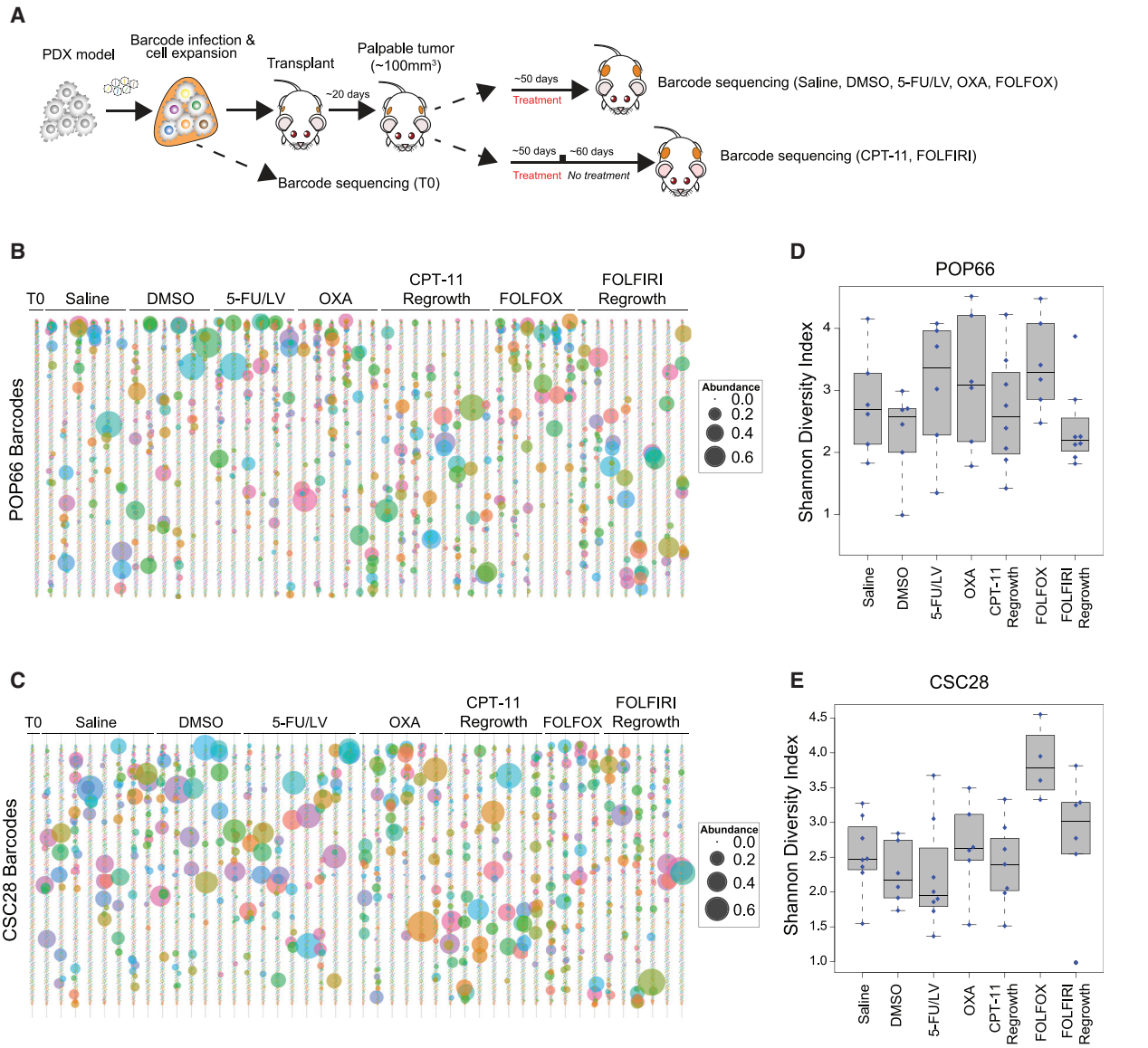


Figure 2. Barcode Complexity in PDXs Post-chemotherapy

(A) Barcoding and *in vivo* treatment schematic for the POP66 and CSC28 CRC PDX models from Figures 1 and S1.

(B and C) Clonal composition of tumors grown in mice from barcoded POP66 (B) or CSC28 (C) models from Figures 1 and S1 following control treatment (saline, DMSO) or chemotherapy. Barcodes are arranged along the y axis according to their starting abundance (T0). Each barcode is assigned a random color for plotting. Only barcodes enriched to at least 1% in one or more tumors are shown.

(D and E) Shannon diversity index for the clones observed in the top 98% of reads for tumors in each treatment group for POP66 (D) and CSC28 (E). No significant differences were found in barcode composition by pairwise Wilcoxon rank-sum test with Holm multiple testing correction.

See also Figure S4 and Table S2.

treatment groups, including treatment regimes that induced a DTP state (CPT-11, FOLFIRI) and those that did not induce a DTP state (5-FU/LV, OXA, FOLFOX). We hypothesized that tumors that went through the DTP state would demonstrate a decrease in barcode complexity and designed the experimental strategy so that we could determine whether the DTP subpopulation represented a pre-existing or *de novo* acquired cellular subset. Unexpectedly, no significant loss in barcode complexity was observed in CRC cells that went through the DTP state *in vivo* (Figures 2B–2E). Notably, the enriched barcodes were unique across all tumors (Figures 2B and 2C), indicating that there was no selection of a pre-existing cell subpopulation that gave rise to DTPs. Treatment with OXA and FOLFOX, which did not induce a DTP state, did not show any change in barcode complexity either. There were no significant differences in tumor weights at the time of sacrifice across all treatment groups,

indicating no sequencing bias from the tumors (Figures S4C and S4D). Although there were significant differences in the time (days) when tumors grew in mice (Figures S4E and S4F), as expected considering the CPT-11 and FOLFIRI tumors went through a DTP state, these differences were not associated with any significant difference in the total number of barcodes detected (Table S2), barcode complexity (Figures 2B–2E), or maximum clone size (Figures S4G and S4H). POP66 CPT-11-resistant tumors were the only tumors to demonstrate a significant decrease in clonal complexity (Figures S4I–S4K). Collectively, our findings show no loss of clonal complexity based on barcode composition of tumor cells that emerged from a DTP state, indicating that DTPs do not represent a cell subpopulation in these CRC models.

An Equipotent Model of Clonal Dynamics in Response to Chemotherapy

The results from the barcoding experiments were unexpected because, based on the significant decrease in tumor size during treatment, we expected a decrease in barcode complexity after tumors regrew following the DTP state. The lack of correlation in the identity of enriched barcodes among biological replicates of the experiment suggested that clonal dominance (evidenced by barcode enrichment) is not primarily driven by pre-existing selective advantages (Shakiba et al., 2019). However, it remained unclear whether the observed variability in clonal output was due to some clones developing selective growth advantages. To address this, we first analyzed clone size distributions and observed that the barcode distributions remained largely ''invariant'' across the two control groups as well as the different chemotherapy treatment groups, including tumors that went through a DTP state (Goyal et al., 2015; Shakiba et al., 2019). This invariance is shown visually by the overlapping cumulative fraction of clones (Figures 3A, 3B, S5A, and S5B). The similarity in clonal dynamics between the control and treated tumors led us to hypothesize that the variability in clonal output is not primarily driven by developing fitness differences. This led us to question whether the observed barcode heterogeneity emerges from stochastic birth-death cellular dynamics where all cells are equipotent (i.e., no fitness differences across the population). To address this, we used more comprehensive mathematical modeling with clonal resolution.

Our mathematical model for clone size distribution accurately captured the experimentally observed clonal dynamics. A striking feature of the experimental clone sizes was the linear relationship between the cumulative fraction of clones and the relative clone size on a log-log plot. We found that this log-linear relationship (Yule, 1925), quantified as the slope on the log-log plot, was consistently between a tight range of 0.3–0.5 across all 80 tumors (Figures 3C and 3D). Our model demonstrated that this linear relationship naturally emerged from a proliferative hierarchy and subexponential (also referred to as logistic or power law) tumor growth kinetics (Gaimann et al., 2020), which has been widely observed and analyzed previously across many different tumors (West and Newton, 2019). In our model, injected cancer cells first exit the slow-cycling state into a fast-cycling state to form the tumor. The variability in clonal output in our model was primarily driven by the stochastic latency in exit from the slow-cycling state, while the slope depends on the tumor growth kinetics. In particular, if $\dot{N} \sim N^{1-\alpha}$, where N is the size of the tumor and α quantifies the deviation from exponential growth (West and Newton, 2019; Figure 3E; note that $\alpha = 0$ corresponds to exponential growth), the clone size distribution is $p(n) \sim n^{-(1+\alpha)}$, resulting in cumulative distribution $Q(n) = \sum_{k>n} p(k) \sim n^{-\alpha}$ (Figure 3F; see STAR Methods for model details). Our mathematical model demonstrates that tumor cells possess an equipotent capacity to form DTPs. Notably, the equipotency extends to the ability of cancer cells to enter the DTP state as well as their ability to exit the DTP state in a stochastic fashion.

DTP-State Cells Express an Embryonic Diapause Gene Signature

We next performed RNA sequencing (RNA-seq) to gain insight into the mechanisms that underlie the CPT-11-induced DTP state (Table S3). The data revealed that DTPs are clearly distinct from vehicle-treated controls (Figures 4A–4C and S6A), with 1,165 genes upregulated and 928 genes downregulated in the DTPs (Figure 4B; Table S4). Genes with roles in metabolism, cell cycle regulation, translation, and RNA processing are highly enriched among genes downregulated in DTPs (Figures 4D and S6B; Table S5). Interestingly, cells that exited the DTP state and led to tumor regrowth are indistinguishable from vehicle-treated controls (Figures 4A–4C and S6A). DTPs are also transcriptionally distinct from tumor cells that acquired resistance to CPT-11 upon continued long-term treatment and eventually gave rise to tumors while on CPT-11 (Figures S6C–S6E; Table S5). Thus, DTP cells represent a reversible state induced by drug treatment but unrelated to classic irreversible drug resistance.

The equipotent capacity of tumor cells to enter the DTP state for survival prompted us to examine whether the tumor cells were employing an evolutionary conserved mechanism to survive the environmental stress associated with chemotherapy exposure. Furthermore, the concerted downregulation of pathways (Myc and mTORc1) and crucial cellular growth processes (cell cycle, translation, and RNA processing) in DTPs recapitulated embryonic developmental pausing, also known as diapause (Figures 4D and S6B). It has been shown previously that mammalian target of rapamycin (mTOR) inhibition induces a reversible suspended pluripotent state in mouse blastocysts and embryonic stem cells (ESCs) (Bulut-Karslioglu et al., 2016). This paused state mimics diapause, a natural phenomenon of transient developmental arrest upon adverse environmental conditions (Fenelon et al., 2014). Therefore, we compared the transcriptomes of vehicle control, DTP-state, and CPT-11 regrowth tumors to previously published gene expression data of paused ESCs and *in vivo* diapaused embryos (STAR Methods; Table S4) and found that DTPs share significant similarities with both models (Figures 4E, 4F, and S6E). We next compared the transcriptomes of vehicle control and DTP-state tumors with previously published gene expression data of various embryonic developmental stages, including diapause (Boroviak et al., 2015). This analysis shows extensive similarities between CPT-11 DTP-state tumors and diapaused epiblasts but not other embryonic developmental stages (Figure S6F). To further explore this relationship, we combined expression data from *in vitro*-paused ESCs with *in vivo* embryonic diapause to develop a

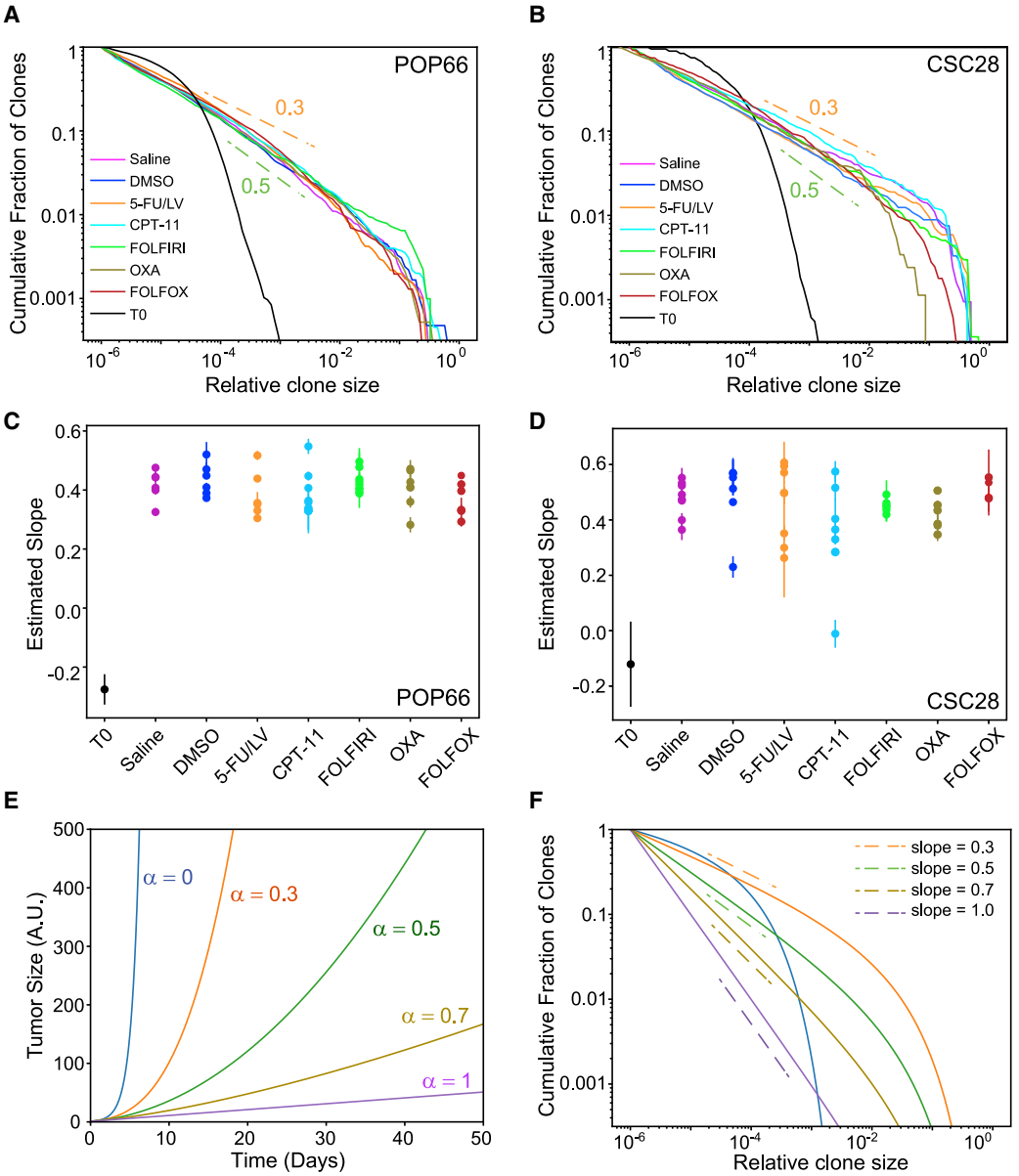


Figure 3. Clonal Composition Remains Invariant across Treatment

(A and B) Mean cumulative clone size distribution for POP66 (A) and CSC28 (B) PDXs across treatments.

(C and D) Estimated power law slope for individual tumors across treatments for POP66 (C) and CSC28 (D) PDXs.

(E) Tumor growth kinetics as a function of interaction parameter (α).

(F) Dependence of cumulative clone size distribution as a function of α.

See also Figure S5.

gene expression signature characteristic of embryonic pausing (STAR Methods; Table S6; Boroviak et al., 2015; Bulut-Karslioglu et al., 2016). Compared with vehicle control or CPT-11 regrowth tumors, DTP-state tumors have the highest embryonic pausing signature score, indicating that changes in their transcriptome resemble that of paused ESCs and embryonic diapause (Figures 4G, S6G, and S6H). These similarities were further corroborated by analyses of expression signatures of specific signaling pathways associated with embryonic diapause, where DTPs expressed significantly reduced levels of mTOR and Myc response modules (Figures 4D, S6I, and S6J; Bulut-Karslioglu et al., 2016; Scognamiglio et al., 2016). Conversely, DTPs also express increased levels of diapause-associated autophagy genes (Figure S6K; Bulut-Karslioglu et al., 2016; He et al., 2019; Vera-Ramirez et al., 2018). Together, these data reveal that CPT-11-induced DTPs are in a reversible state with extensive

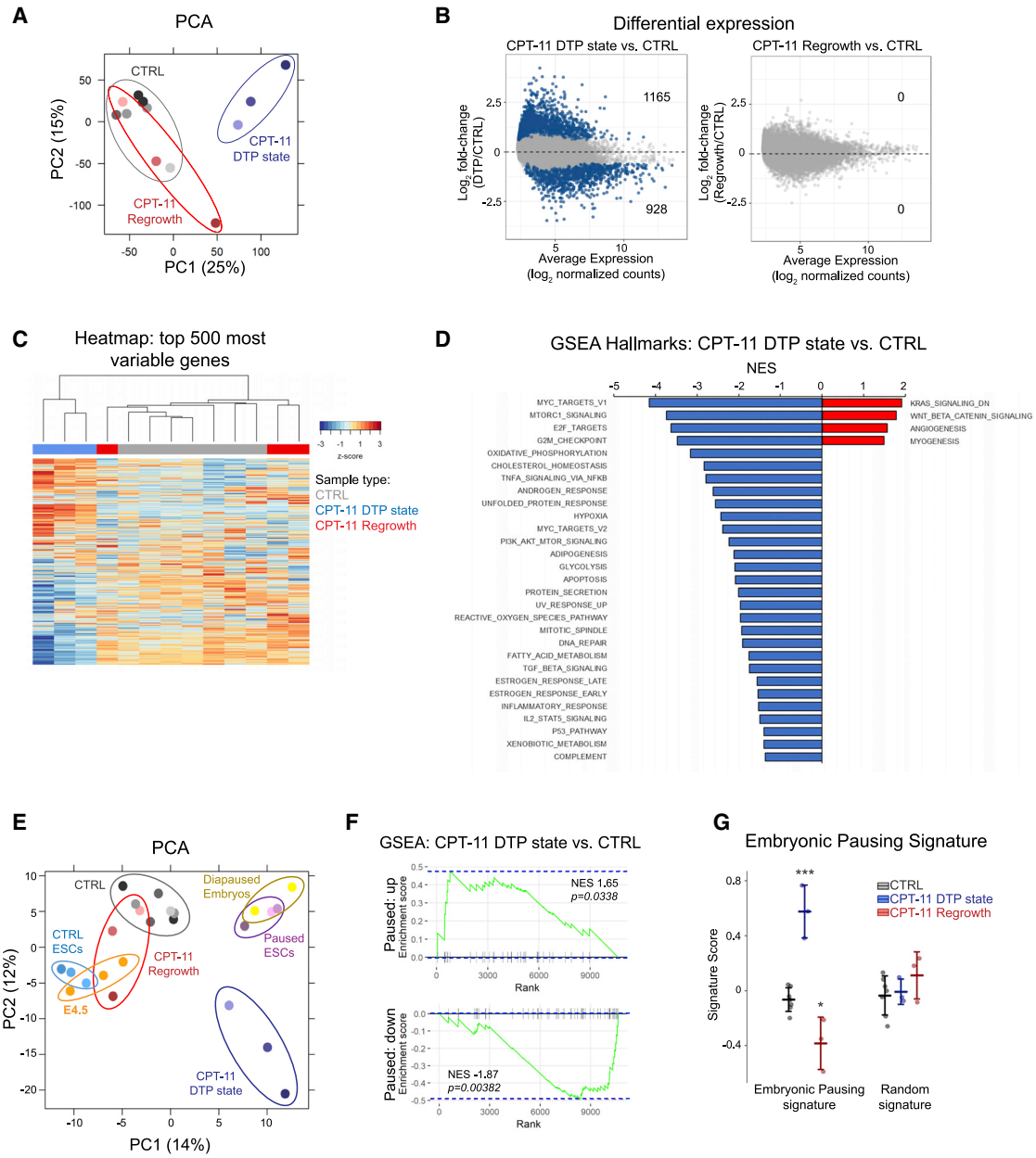


Figure 4. DTP-State Cells Express an Embryonic Pausing Signature

(A) Principal-component analysis (PCA) plot for all expressed genes across control (CTRL), CPT-11 DTP-state, and CPT-11 regrowth tumor samples.

(B) MA plot showing $\log_2$ fold changes in the expression of each gene in CPT-11 DTP-state (left) or CPT-11 regrowth (right) over CTRL samples. Genes differentially expressed (false discovery rate [FDR] < 0.05, fold-change > 1.5 or < 2/3) are shown in blue. Numbers correspond to upregulated (top) and downregulated (bottom) genes.

(C) Heatmap of the top 500 variable genes, with CTRL, CPT-11 DTP-state and CPT-11 Regrowth samples grouped by unsupervised hierarchical clustering.

(D) GSEA of Hallmark pathways measured by normalized enrichment score (NES) for all significantly altered pathways (FDR < 0.05) in DTP tumors.

(E) PCA plot for all expressed genes across tumor samples (CTRL, CPT-11 DTP-state, CPT-11 regrowth), ESCs (CTRL ESCs and paused ESCs), and embryos (embryonic day 4.5 [E4.5] and diapaused embryos).

(F) GSEA of CPT-11 DTP-state tumor samples for genes upregulated ($t > 10$, top panel) or downregulated ($t < -10$, bottom panel) in paused ESCs (versus CTRL ESCs). Genes were pre-ranked by t values for CPT-11 DTP state (versus CTRL). NES and Bonferroni correction adjusted p values are indicated.

(G) Scores of the embryonic pausing signature (n = 124 genes) in CTRL, CPT-11 DTP-state, and CPT-11 regrowth samples. A signature of 124 randomly selected genes is shown as a control random signature. Data are mean ± SD, Welch two-sample t test. *p < 0.05, ***p < 0.001.

See also Figure S6 and Tables S3, S4, S5, and S6.

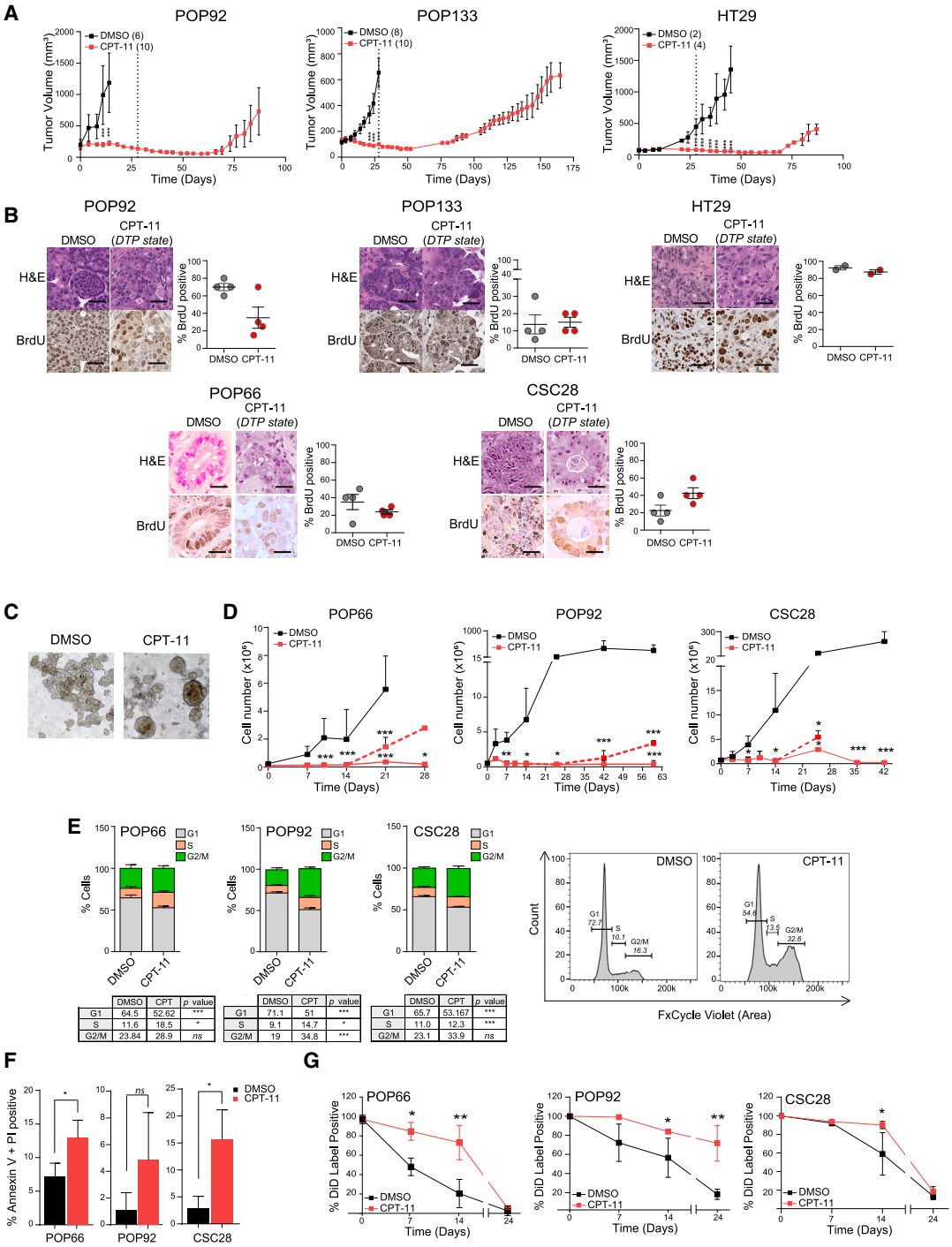


Figure 5. CPT-11-Induced DTP State CRC Xenografts and Cultures Are Slow Cycling

(A) Tumor growth curves of CRC PDXs upon treatment with DMSO or CPT-11. Each point represents mean tumor volume ± SEM. Numbers in parentheses denote the biological replicates in that group.

(B) H&E and BrdU immunohistochemistry staining with quantification of POP92, POP133, and HT29 tumors (from A) and POP66 and CSC28 tumors (from Figure 1) treated with DMSO or CPT-11 for 28 days (20×). Images are representative of 4 (POP92, POP133, POP66, and CSC28) or 2 (HT29) biological replicates. Differences in BrdU incorporation were not statistically significant; unpaired t test with Welch's correction.

(C) Representative bright-field images for POP66 cultures treated with DMSO or CPT-11 for 10 days (20×) *in vitro*.

(legend continued on next page)

transcriptional similarities to embryonic diapause, hereafter referred to as a diapause-like DTP state.

CRC Tumors in the CPT-11-Induced Diapause-like DTP State Are Slow Cycling

Therapeutic pressure has been shown to drive cancer cells into a DTP state to survive the insult by entering a low proliferative state (Hangauer et al., 2017; Hata et al., 2016; Risom et al., 2018; Sharma et al., 2010). Similarly, a defining characteristic of diapause is that cells enter a slow-cycling state (Bulut-Karslioglu et al., 2016; Hand et al., 2016; Kostál et al., 2009; Scognamiglio et al., 2016; Ślusarczyk et al., 2019). Although we observed low Ki-67 expression for tumors in the diapause-like DTP state (Figures 1F and 1G), we performed *in vivo* bromodeoxyuridine (BrdU) labeling to determine the percentage of cancer cells that were cycling while in the DTP state. We first established three additional CRC PDX models of the *in vivo* CPT-11-induced DTP state (Figures 5A, S7A, and S7B; Table S1). A subset of mice were administered BrdU during weeks 3 and 4 of CPT-11 treatment, and tumors were harvested at the end of treatment (4 weeks). BrdU incorporation analysis revealed that almost all tumor cells were positive, with no significant difference in percentage of BrdU positivity between DMSO- and CPT-11-treated tumors (Figure 5B), establishing that the vast majority of cells in the DTP state *in vivo* are cycling, albeit slowly.

In Vitro Characterization of CRC cells in the CPT-11-Induced Diapause-Like DTP State

We next explored the biology of the DTP state *in vitro* using CRC models treated with CPT-11 at 1 μM, a concentration found to cause 50% growth inhibition (GI_{50}). POP66, POP92, and CSC28 cultures treated with CPT-11 demonstrated significant growth suppression but resumed growth upon drug withdrawal on day 14 (Figures 5C and 5D), indicating that CRC cultures also enter a reversible DTP state upon CPT-11 treatment *in vitro*. Consistent with previous reports of the effect of CPT-11 on tumor cells (Kaku et al., 2015; Wang et al., 2018), we found that CPT-11 induced a stall in S and G2/M phases of the cell cycle and a decrease in the G1 fraction (Figure 5E). However, although cell cycle progression was slightly altered, gene set enrichment analysis (GSEA) Hallmarks analysis shows that the transcriptomes of tumors in the diapause-like DTP state, diapaused embryos, and paused ESCs are very distinct from the published transcriptomic profiles of various cell cycle arrest models (Figure S7C). In addition, western blot analysis showed no induction of cyclin-dependent kinase inhibitors in POP66 and POP92 DTP-state cells (Figure S7D). We observed a small but statistically significant increase in apoptosis for POP66 and CSC28 in response to CPT-11 treatment but not for POP92 (Figure 5F). Label retention experiments (Quayle et al., 2018) revealed dilution of the Vybrant DiD dye at a much slower rate in CPT-11-treated cultures compared with the control (Figures 5G and S7E). Additionally, CPT-11-treated cultures exhibited a significant decrease in total RNA abundance (Figure S7F), consistent with the global suppression of transcription seen in embryonic diapause (Bulut-Karslioglu et al., 2016). These data suggest that, similar to our *in vivo* BrdU results as well as diapaused blastocysts (Bulut-Karslioglu et al., 2016), *in vitro* CPT-11-treated cultures in the DTP state are slow cycling and enter a diapause-like state to survive chemotherapeutic stress.

The CPT-11-Induced Diapause-like DTP State Is Maintained via the Autophagy Pathway

Guided by our RNA-seq analysis of tumors in the diapause-like DTP state, we sought to identify specific pathways that drive DTP formation. Diapaused blastocysts are known to downregulate cellular transcription and translation programs via key pathways such as mTOR (Bulut-Karslioglu et al., 2016; Scognamiglio et al., 2016). In line with RNA-seq analysis of tumors in the diapause-like DTP state that revealed significant downregulation of the mTOR pathway (Figure S6I), western blot analysis of DTP-state cultures also showed a decrease in mTOR signaling (Figure 6A). Because previous reports have shown that the Ras/mitogen-activated protein kinase (MAPK) pathway is activated upon mTOR inhibition (Carracedo et al., 2008; Zhang et al., 2017), we also observed extracellular signal-regulated kinase 1/2 (ERK 1/2) activation in DTP-state cultures by western blot analysis (Figure 6A). Furthermore, similar to CPT-11-treated cells, POP92 and CSC28 cultures treated with the mTORC1/2 inhibitor INK 128 (GI_{50}, 25 nM) showed significant growth suppression but no increase in apoptosis compared with DMSO-treated controls (Figures 6B and 6C; Hsieh et al., 2012). The cultures resumed growth upon withdrawal of the mTOR inhibitor on day 14 (Figure 6B). These data indicate that, similar to CPT-11 treatment, CRC cells enter a reversible DTP state upon mTOR inhibition. This finding supports the functional similarities between mTOR inhibition-dependent diapaused blastocysts and paused ESCs and the CRC diapause-like DTP state we identified (Bulut-Karslioglu et al., 2016).

mTOR is known to phosphorylate and inactivate autophagy regulators (Chan, 2009). In the context of diapaused blastocysts, mTOR downregulation results in upregulation of key autophagy genes, including ULK1/ATG1 (Bulut-Karslioglu et al., 2016). Furthermore, induction of autophagy plays a key functional role in mediating diapause because blastocysts treated with an autophagy inhibitor (the ULK1 inhibitor SBI-0206965) exit the

(D) Cell growth curves of POP66, POP92, and CSC28 CRC cultures treated with DMSO (black) or CPT-11 (1 μM; red) *in vitro*. A solid line indicates cell growth on treatment, and a dashed line indicates cell growth when treatment was stopped on day 14. Data are mean ± SD, n = 4 independent experiments, two-way ANOVA.

(E) Cell cycle analysis of CRC cultures treated with DMSO or CPT-11 for 14 days. Data are mean ± SEM, n = 3 independent experiments, two-way ANOVA. Representative flow plots are shown.

(F) Apoptosis/necrosis analysis (Annexin V/PI label) on CRC cultures treated with DMSO or CPT-11 for 5 days. Data are mean ± SD; at least 3 independent experiments; t test.

(G) Cytofluorimetric analysis of percent Vybrant DiD-labeled positive cultures treated with DMSO or CPT-11 *in vitro*. Treatment was stopped on day 14. Data are mean ± SEM; n = 3 independent experiments; t test.

*p < 0.05, **p < 0.01, ***p < 0.001; ns, not significant. See also Figure S7.

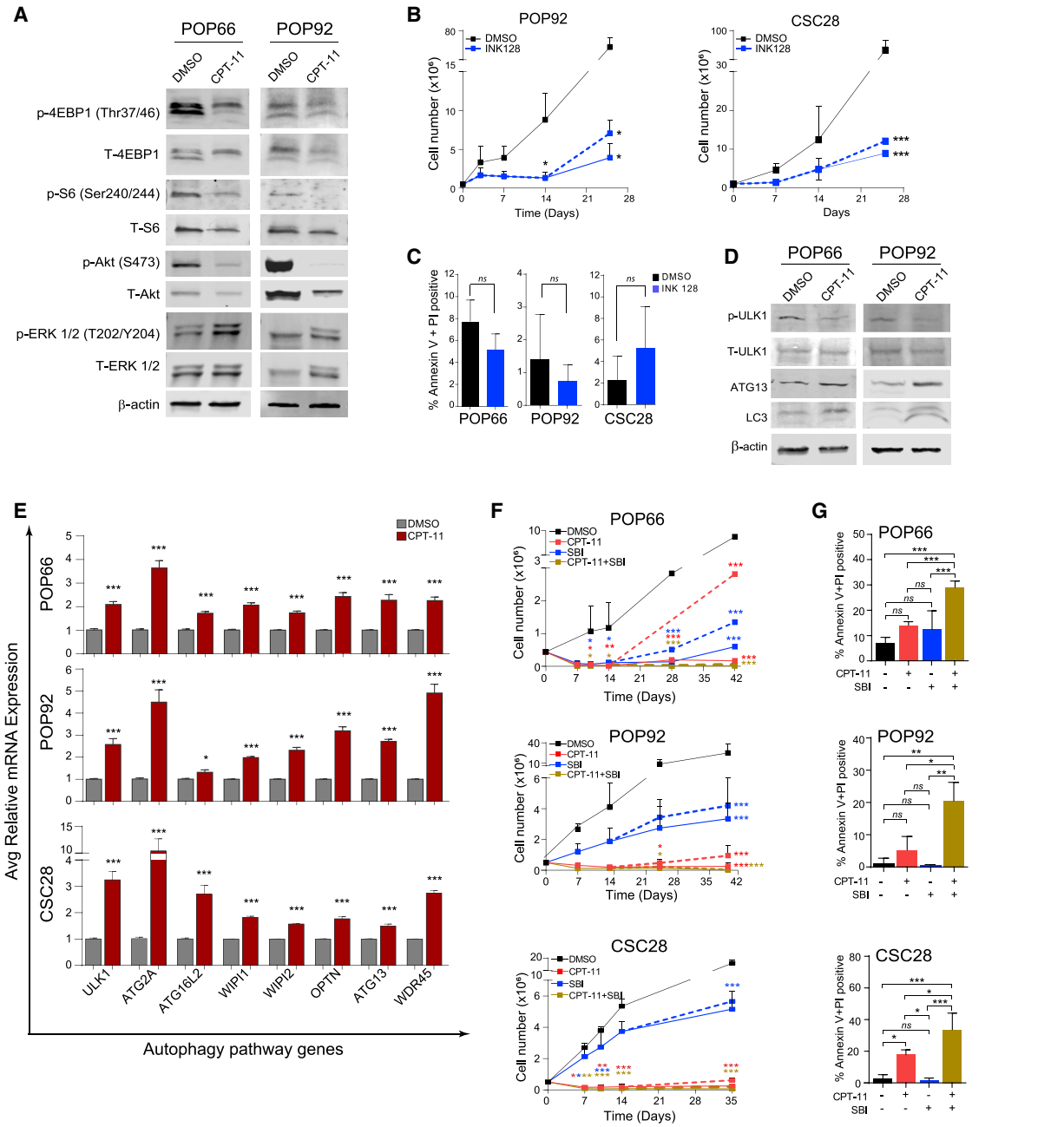


Figure 6. The CPT-11-Induced Diapause-like DTP State Is Maintained by Upregulation of Autophagy

(A) Western blot analysis of the mTOR pathway in POP66 and POP92 cultures treated with DMSO or CPT-11 (1 μM) for 14 days *in vitro* (representative of at least 3 independent experiments).

(B) Cell growth curves of CRC cultures treated with DMSO (black) or the mTOR inhibitor INK 128 (25 nM, blue). Solid lines indicate ongoing treatment, and dashed lines indicate cell growth after treatment was stopped on day 14 day. Data are mean ± SD; n = 3 independent experiments; two-way ANOVA.

(C) Apoptosis/necrosis analysis (Annexin V/PI label) on CRC cultures treated with DMSO or INK 128 for 5 days. Data are mean ± SD; n = 3 independent experiments; one-way ANOVA.

(D) Western blot analysis of autophagy pathway proteins in CRC cultures treated with DMSO or CPT-11 for 14 days *in vitro* (representative of at least 3 independent experiments).

(E) qRT-PCR analysis of autophagy pathway genes in CRC cultures treated with CPT-11 for 7 days *in vitro*. Values are relative to DMSO, normalized to RPLP0, mean ± SEM; n = 3 independent experiments; t test.

(legend continued on next page)

diapaused state (Bulut-Karslioglu et al., 2016; Lee et al., 2011). Consistent with these published findings, RNA-seq analysis of tumors in the diapause-like DTP state demonstrated upregulation of key autophagy genes (Figure S6K). These results were further confirmed *in vitro*, where diapause-like DTP cultures showed a decrease in inhibitory phosphorylated ULK1 and an increase in ATG13 and LC3 protein expression (Figure 6D) in addition to increased gene expression of multiple autophagy pathway mediators (Figure 6E).

We identified ULK1 as a targetable protein in the pathway and treated cells with SBI-0206965 (SBI; GI_{50}, 2.5 μM) (Egan et al., 2015). Although all CRC cultures exhibited significant growth delay in response to SBI treatment alone, the cells resumed proliferation when SBI treatment was discontinued (Figure 6F). Cultures treated with CPT-11 responded as described above, with induction of the diapause-like DTP state followed by recovery when CPT-11 treatment was discontinued. In contrast, cells treated with CPT-11 and SBI combination therapy resulted in significant growth suppression with negligible recovery when treatment was stopped. Combination treatment resulted in significant induction of apoptosis compared with DMSO, CPT-11, or SBI treatment alone (Figure 6G). Furthermore, combination treatment of cells with CPT-11 and chloroquine (CQ; 20 μM), a general autophagy inhibitor, resulted in robust induction of apoptosis compared with either treatment alone (Figure S7G). These data show that, similar to diapaused blastocysts, the diapause-like DTP state in our CRC models is dependent on upregulation of the autophagy program and that targeting this pathway with an ULK1 inhibitor prevents its manifestation.

Clinical Relevance of the Embryonic Pausing Signature

We sought to investigate whether the embryonic pausing signature that defines our tumors in the diapause-like DTP state can also define what is noted in literature as minimal residual disease (MRD) in cancer. Interestingly, MRD transcriptomes derived from animal models and individuals across a range of tumor types (acute lymphoblastic leukemia, prostate cancer, and melanoma) significantly correlated with the embryonic pausing signature, similar to tumors in the diapause-like DTP state (Figure 7A; Ebinger et al., 2016; Rambow et al., 2018; Sowalsky et al., 2018). This suggests that our embryonic pausing signature can identify MRD.

Given the potential role of DTPs in mediating therapeutic outcome and tumor relapse, we investigated whether our signature, which defined tumors in the diapause-like DTP state, could provide prognostic stratification in the TCGA (The Cancer Genome Atlas) cohort of individuals with CRC (Cancer Genome Atlas, 2012). The samples from individuals with CRC were divided into two groups, defined as having a "high" or "low" embryonic pausing signature. Kaplan-Meier analysis showed that individuals with a high signature score had significantly worse overall survival than individuals with a low signature score (Figure 7B). Furthermore, we analyzed this CRC TCGA dataset for the autophagy signature (Gene Ontology [GO]: 0010506). Kaplan-Meier analysis showed that individuals with CRC who had a high autophagy score had significantly worse overall survival than those with a low autophagy score (Figure 7C). We extended our analysis to various cancers of the TCGA and found that, for the indicated cancer cohorts, a high autophagy signature corresponds to poor overall survival ($p < 0.05$; hazard ratio [HR]; Figure 7D). These data suggest that patients with a high diapause or autophagy score have tumor cells that would be refractory to therapy. We propose that this signature provides a way to identify individuals with refractory disease and uncovers innovative therapeutic intervention strategies to target the diapause-like DTP state responsible for relapse.

DISCUSSION

In this study, we leveraged cellular barcoding in two CRC PDX models to show that there is no loss of genetic or barcode heterogeneity in cancer cells that go through a chemotherapy-induced DTP state. Instead, our results demonstrate that cancer cells are equipotent in their capacity to enter the DTP state, which they accomplish by activating the evolutionarily conserved survival strategy of diapause. Similar to diapause, the diapause-like DTP state is characterized by a decrease in mTOR pathway activity and an increase in autophagy (Bulut-Karslioglu et al., 2016). Combination treatment of chemotherapy and autophagy inhibitors induced DTP cell death, providing insight into therapeutic opportunities to target cancer cells in the diapause-like DTP state.

Some studies indicate that DTPs pre-exist in the original cancer cell population and that chemotherapy exposure promotes selection of this subpopulation or that chemotherapy induces a phenotypic transition to a DTP state in a subpopulation of tumor cells (Guler et al., 2017; Liau et al., 2017; Sharma et al., 2010; Roesch et al., 2010). Our experiments were designed so that there would be multiple tumors with comparable starting barcode populations, allowing us to distinguish pre-existing from DTP clones acquired *de novo*. If a pre-existing cancer cell subpopulation was the major source of DTPs, then we would expect to see selective enrichment of the same barcodes in replicate tumors. However, if DTPs were driven by a stochastic process, then we would expect distinct barcoded populations to emerge in independent tumor replicates (Bhang et al., 2015; Merino et al., 2019; Seth et al., 2019). In our models, the same barcodes were not detected across replicate tumors, indicating that DTPs were not a pre-existing subset.

We anticipated that the bottleneck associated with DTP formation in our models would be reflected in loss of barcode

(F) Cell growth curves of CRC cultures treated *in vitro* with DMSO (black), CPT-11 (red), SBI-0206965 (SBI; 2.5 μM, blue) or a combination (CPT-11+SBI, gold). Solid lines indicate cell growth on treatment, and dashed lines indicate cell growth when treatment was stopped on day 14. Data are mean ± SEM; n = 3 independent experiments; two-way ANOVA.

(G) Apoptosis/necrosis analysis (Annexin V/PI label) performed on CRC cultures treated with DMSO, CPT-11, SBI or a combination for 5 days. Data are mean ± SD; n = 3 independent experiments; one-way ANOVA.

*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$, ns = not significant.

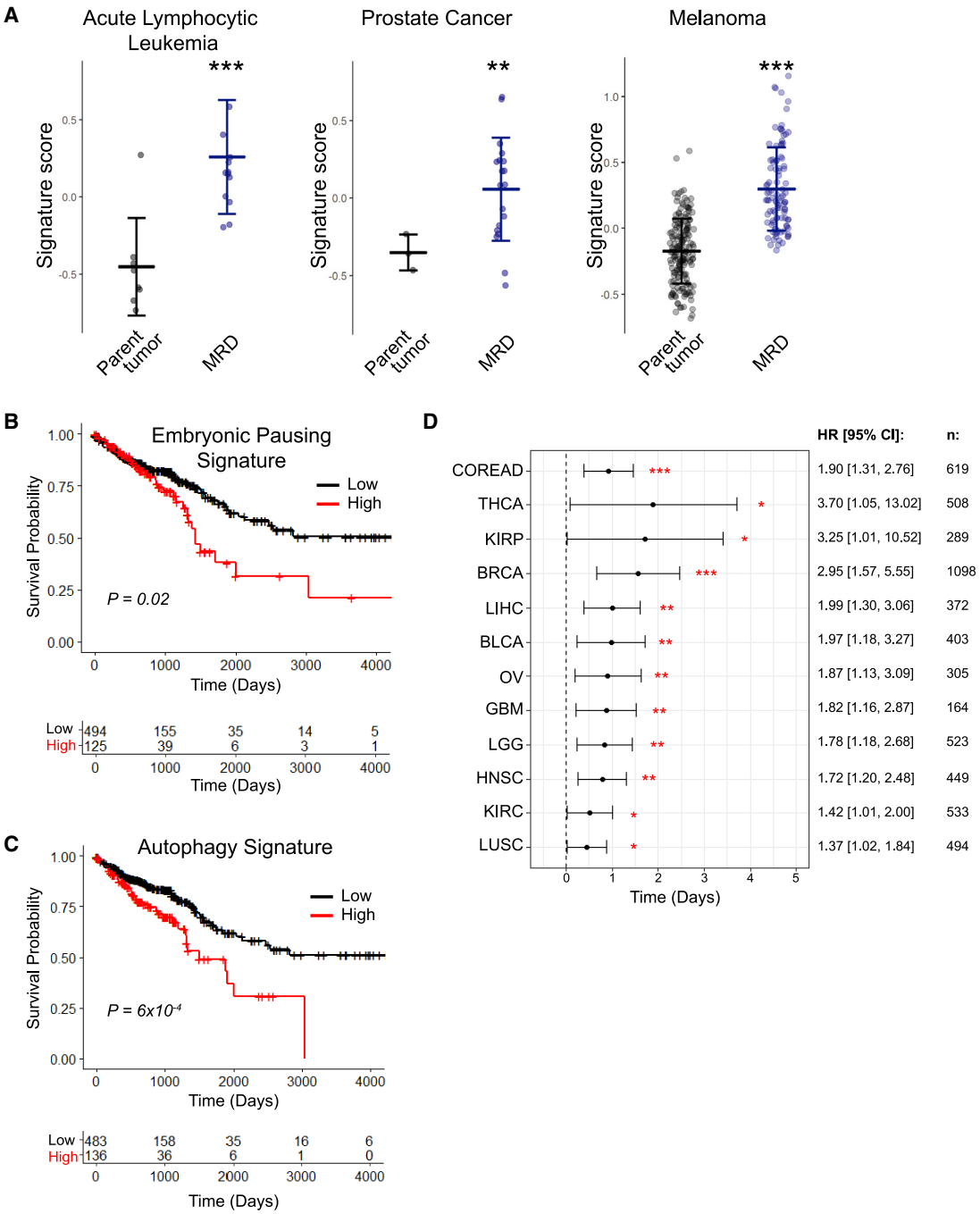


Figure 7. The Embryonic Pausing Signature and Autophagy Signature Applied to Various Cancer Models

(A) Scores of the embryonic pausing signature in previously reported studies related to minimal residual disease (MRD; GEO: GSE83142, GSE102124, and GSE116237); Welch two-sample t test, **p < 0.01, ***p < 0.001.

(B and C) Overall Kaplan-Meier plot for TCGA data for individuals with CRC, stratified to high and low groups based on (B) the embryonic pausing signature (p = 0.02, log rank test) or (C) GO autophagy signature (GO: 0010506; $p = 6 \times 10^{-4}$, log rank test).

(D) Forest plot showing the overall survival hazard ratio (HR) for the indicated cancers of TCGA based on GO autophagy signature.*p < 0.05, **p < 0.01, ***p < 0.001.

complexity (Guler et al., 2017; Sharma et al., 2010). Unexpectedly, no significant selection was observed based on barcode complexity in tumors that regrew following the DTP state. It could be proposed that the observed barcode variability across tumors is driven by selective dynamics of multiple rapidly mutating clones (Shakiba et al., 2019). However, WES analysis did not

reveal any significant generation of mutations in regrowth tumors, negating this possibility. Instead, quantitative analysis of barcode composition suggested that the observed variability in barcode complexity is explained by a model that combines proliferative hierarchy with sub-exponential tumor growth kinetics (Goyal et al., 2015; Lan et al., 2017; West and Newton, 2019). Here the barcode complexity does not result from genetic diversity based on selective dynamics of cells with fitness differences but results from stochastic latency in the exit times from a slow-cycling state (Lan et al., 2017). Thus, our data strongly support the "equipotent" model, in which all tumor cells have equal capacity to enter and exit the DTP state to survive therapy, refuting the idea that a small subpopulation of cells in the parent tumor are capable of becoming DTPs.

There is increasing interest in approaching therapeutic strategies in cancer by maintaining tumor cells in a DTP state. However, our prolonged treatment of a subset of mice with CPT-11 (5–6 months) suggests that, over time, tumors in the DTP state act as a reservoir for development of genetically resistant clones. Tumors that grew out while on prolonged CPT-11 treatment had decreased genetic diversity, indicating that genetic clonal selection occurred while the cells remained in the prolonged DTP state. Our findings are consistent with those of Ramirez et al. (2016), who demonstrated that persister-derived erlotinib-resistant colonies acquired diverse resistance mechanisms over time, suggesting that the DTP state provides a latent reservoir of cells for emergence of drug-resistant clones. In light of this and our findings, tumor cells should not be held in a persister state. Instead, therapeutic strategies should be developed to eliminate persisters or prevent their formation.

Transcriptional profiling of tumors in the DTP state showed gene expression patterns that closely resembled hormonally and chemically induced diapause. Although the molecular mechanisms governing mouse embryonic diapause remain largely unknown, two recent studies identified downregulation of the Myc and mTOR pathways as a key player in embryonic diapause (Bulut-Karslioglu et al., 2016; Scognamiglio et al., 2016). Inhibition of Myc in mouse ESCs *in vitro* resembled diapause in blastocysts *in vivo*, and inhibition of Myc for 18 h in blastocysts was compatible with development to term (Scognamiglio et al., 2016). Inhibition of mTOR induced reversible pausing of mouse blastocyst development and allowed their prolonged culture *ex vivo* up to 22 days while retaining the ability to develop to term (Bulut-Karslioglu et al., 2016). Interestingly, following Myc depletion, murine blastocysts entered a dormant state, exiting the cell cycle (Scognamiglio et al., 2016). In contrast, the diapause state induced *ex vivo* by mTOR inhibition was not associated with dormancy but a slow-cycling state (Bulut-Karslioglu et al., 2016). Similar to the mTOR inhibitor-induced diapause state, CPT-11-induced DTPs in our study were not arrested and remained engaged in a slow-cycling state for a prolonged time period. Furthermore, naturally diapaused blastocysts *in vivo* and paused blastocysts *ex vivo* displayed pronounced reductions in mTOR activity, alterations in chromatin modifications, hypotranscription, and activation of the key autophagy regulator ULK1 (Bulut-Karslioglu et al., 2016). Autophagy was shown to play an important functional role because paused blastocysts co-treated with autophagy inhibitors (the ULK1 inhibitor SBI or CQ) resulted in decreased survival of the paused state. Here, downregulation of the Myc and mTOR modules as well as upregulation of key autophagy genes were observed when tumors were in the diapause-like DTP state. Furthermore, co-treatment with autophagy inhibitors and CPT-11 significantly inhibited acquisition of a DTP state and resulted in robust cell death. Thus, DTPs not only transcriptionally recapitulate diapause but also functionally depend on pathways that drive embryonic viability during diapause.

The understanding of DTPs is at a relatively nascent stage, and whether DTPs differ depending on specific chemotherapies or targeted agents is currently not established. Some of the key driver genes and pathways identified in DTP models to date include GPX4, LINE-1 repression, IGF-1R, and oxidative phosphorylation (Echeverria et al., 2019; Guler et al., 2017; Hangauer et al., 2017; Sharma et al., 2010). However, we identified that DTPs activate a developmentally conserved program to survive the harsh environment created by chemotherapy, and understanding this mechanism will be key to devising methods to target this state. Future studies of the diapause-like DTP state in the context of cancer should include developing robust PDX preclinical DTP models in response to standard-of-care chemotherapy and targeted agents as well as syngeneic models that expand our understanding of how the diapause-like DTP state is influenced by the immune system. Targeting the diapause-like DTP state will be essential for developing therapeutic strategies designed to eliminate cancer cells when they are potentially in their most vulnerable state.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead Contact
  - Materials Availability
  - Data and Code Availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Mice
  - Human tissue and patient-derived xenografts
  - Cell lines and Primary Cell Cultures
- METHOD DETAILS
  - DNA barcode library
  - Lentivirus production, infection, cell sorting
  - Chemicals
  - *In vivo* chemotherapy drug treatments
  - BrdU administration
  - Tissue collection and histological analysis
  - Digestion of xenograft tumor tissue
  - DNA extraction for barcode sequencing
  - DNA barcode amplification and sequencing
  - Barcode processing and analysis
  - Exome sequencing and analysis
  - Pairtree Intratumoral Heterogeneity Analysis
  - Purification of tumor cells from DTP-state xenografts
  - RNA extraction

- RNA-sequencing and data processing
- RNA-seq bioinformatic analysis
- Survival analysis
- *In vitro* chemotherapy drug treatments
- Cell growth analysis
- Fluorescent dye label retention assay
- Real-time PCR
- Western blot analysis
- Cell cycle and apoptosis/necrosis
- Mathematical modeling, stochastic simulations

● QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at https://doi.org/10.1016/j.cell.2020.11.018.

ACKNOWLEDGMENTS

We thank Aaron Schimmer, Eric Lechman, Faiyaz Notta, Steven Gallinger, Laurie Ailles, John Dick, Daniel De Carvalho, James Kennedy, and Michael Taylor for critical review of this manuscript; Ryan Dowling for technical advice; the Pathology Research Program for assistance with immunohistochemistry; the SickKids-UHN Flow Cytometry Facility for assistance with fluorescence-activated cell sorting; the UHN Animal Resources Centre for animal housing and veterinary support; The Center for Applied Genomics for sequencing support; and the UHN BioBank for providing CRC samples. This work was supported by funds from the Canadian Institutes of Health Research (FDN-148479 to C.A.O., MOP-142375 to J.M., and project grant 420231 to M.R.-S.) and the Terry Fox Foundation (TFRI-1029 to C.A.O.). M.R.-S. is supported by a Canada 150 Research Chair in Developmental Epigenetics. S.G. is supported by the NSERC Discovery Grants Program (RGPIN 1505). C.A.O is supported by a Canada Research Chair in Translational Colorectal Cancer Research.

AUTHOR CONTRIBUTIONS

Conceptualization, C.A.O.; Methodology and Validation, J.H., S.K.R., A.M.L.N., Y.W., N.M.P., J.M., M.R.-S., S.G., C.A.O.; Investigation, S.K.R., J.H., A.M.L.N., E.B.L.L., Y.W., E.L.-F., C.L., and F.S.; Formal Analysis, S.K.R., J.H., E.C., A.M.L.N., K.R.B., S.M., S.G., J.P.B., J.A.W., A.S.M., A.P., and H.H.H.; Writing – Original Draft, S.K.R., J.H., E.C., K.R.B., A.S.M., J.P.B., S.G., J.M., M.R.-S., T.J.P., B.H.-K., and C.A.O.; Writing – Review and Editing, S.K.R., J.H., E.C., K.R.B., A.S.M., J.P.B., J.A.W., Q.M., S.G., J.M., M.R.-S., T.J.P., B.H.-K., and C.A.O.; Visualization, S.K.R., J.H., E.C., K.R.B., S.G., and A.S.M.; Supervision, S.G., J.M., B.H.-K., T.J.P., Q.M., M.R.-S., and C.A.O.

DECLARATION OF INTERESTS

J.M. is a shareholder in Northern Biologics and Pionyr Immunotherapeutics and an advisor for Century Therapeutics and Aelian Biotechnology.

Received: February 7, 2020
Revised: August 25, 2020
Accepted: November 10, 2020
Published: January 7, 2021

REFERENCES

Balaban, N.Q., Merrin, J., Chait, R., Kowalik, L., and Leibler, S. (2004). Bacterial persistence as a phenotypic switch. Science *305*, 1622–1625.

Balaban, N.Q., Gerdes, K., Lewis, K., and McKinney, J.D. (2013). A problem of persistence: still more questions than answers? Nat. Rev. Microbiol. *11*, 587–591.

Benjamin, D., Sato, T., Cibulskis, K., Getz, G., Stewart, C., and Lichtenstein, L. (2019). Calling Somatic SNVs and Indels with Mutect2. bioRxiv. https://doi.org/10.1101/861054.

Bhang, H.E., Ruddy, D.A., Krishnamurthy Radhakrishna, V., Caushi, J.X., Zhao, R., Hims, M.M., Singh, A.P., Kao, I., Rakiec, D., Shaw, P., et al. (2015). Studying clonal dynamics in response to cancer therapy using high-complexity barcoding. Nat. Med. *21*, 440–448.

Boroviak, T., Loos, R., Lombard, P., Okahara, J., Behr, R., Sasaki, E., Nichols, J., Smith, A., and Bertone, P. (2015). Lineage-Specific Profiling Delineates the Emergence and Progression of Naive Pluripotency in Mammalian Embryogenesis. Dev. Cell *35*, 366–382.

Bowen, M.E., McClendon, J., Long, H.K., Sorayya, A., Van Nostrand, J.L., Wysocka, J., and Attardi, L.D. (2019). The Spatiotemporal Pattern and Intensity of p53 Activation Dictates Phenotypic Diversity in p53-Driven Developmental Syndromes. Dev. Cell *50*, 212–228.e6.

Brauner, A., Fridman, O., Gefen, O., and Balaban, N.Q. (2016). Distinguishing between resistance, tolerance and persistence to antibiotic treatment. Nat. Rev. Microbiol. *14*, 320–330.

Bulut-Karslioglu, A., Biechele, S., Jin, H., Macrae, T.A., Hejna, M., Gertsenstein, M., Song, J.S., and Ramalho-Santos, M. (2016). Inhibition of mTOR induces a paused pluripotent state. Nature *540*, 119–123.

Cancer Genome Atlas, N.; Cancer Genome Atlas Network (2012). Comprehensive molecular characterization of human colon and rectal cancer. Nature *487*, 330–337.

Carracedo, A., Ma, L., Teruya-Feldstein, J., Rojo, F., Salmena, L., Alimonti, A., Egia, A., Sasaki, A.T., Thomas, G., Kozma, S.C., et al. (2008). Inhibition of mTORC1 leads to MAPK pathway activation through a PI3K-dependent feedback loop in human cancer. J. Clin. Invest. *118*, 3065–3074.

Cerami, E., Gao, J., Dogrusoz, U., Gross, B.E., Sumer, S.O., Aksoy, B.A., Jacobsen, A., Byrne, C.J., Heuer, M.L., Larsson, E., et al. (2012). The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. Cancer Discov. *2*, 401–404.

Chan, E.Y. (2009). mTORC1 phosphorylates the ULK1-mAtg13-FIP200 autophagy regulatory complex. Sci. Signal. *2*, pe51.

Clauset, A., Shalizi, C.R., and Newman, M.E.J. (2009). Power-Law Distributions in Empirical Data. SIAM Rev. *51*, 661–703.

Conklin, J.F., Baker, J., and Sage, J. (2012). The RB family is required for the self-renewal and survival of human embryonic stem cells. Nat. Commun. *3*, 1244.

Deng, L., Li, C., Chen, L., Liu, Y., Hou, R., and Zhou, X. (2018). Research advances on embryonic diapause in mammals. Anim. Reprod. Sci. *198*, 1–10.

Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics *29*, 15–21.

Dobson, S.M., García-Prat, L., Vanner, R.J., Wintersinger, J., Waanders, E., Gu, Z., McLeod, J., Gan, O.I., Grandal, I., Payne-Turner, D., et al. (2020). Relapse-Fated Latent Diagnosis Subclones in Acute B Lineage Leukemia Are Drug Tolerant and Possess Distinct Metabolic Programs. Cancer Discov. *10*, 568–587.

Ebinger, S., Özdemir, E.Z., Ziegenhain, C., Tiedt, S., Castro Alves, C., Grunert, M., Dworzak, M., Lutz, C., Turati, V.A., Enver, T., et al. (2016). Characterization of Rare, Dormant, and Therapy-Resistant Cells in Acute Lymphoblastic Leukemia. Cancer Cell *30*, 849–862.

Echeverria, G.V., Ge, Z., Seth, S., Zhang, X., Jeter-Jones, S., Zhou, X., Cai, S., Tu, Y., McCoy, A., Peoples, M., et al. (2019). Resistance to neoadjuvant chemotherapy in triple-negative breast cancer mediated by a reversible drug-tolerant state. Sci. Transl. Med. *11*, eaav0936.

Egan, D.F., Chun, M.G., Vamos, M., Zou, H., Rong, J., Miller, C.J., Lou, H.J., Raveendra-Panickar, D., Yang, C.C., Sheffler, D.J., et al. (2015). Small Molecule Inhibition of the Autophagy Kinase ULK1 and Identification of ULK1 Substrates. Mol. Cell *59*, 285–297.

Ewels, P., Magnusson, M., Lundin, S., and Käller, M. (2016). MultiQC: summarize analysis results for multiple tools and samples in a single report. Bioinformatics *32*, 3047–3048.

Favero, F., Joshi, T., Marquard, A.M., Birkbak, N.J., Krzystanek, M., Li, Q., Szallasi, Z., and Eklund, A.C. (2015). Sequenza: allele-specific copy number and mutation profiles from tumor sequencing data. Ann. Oncol. *26*, 64–70.

Fenelon, J.C., and Renfree, M.B. (2018). The history of the discovery of embryonic diapause in mammals. Biol. Reprod. *99*, 242–251.

Fenelon, J.C., Banerjee, A., and Murphy, B.D. (2014). Embryonic diapause: development on hold. Int. J. Dev. Biol. *58*, 163–174.

Gaimann, M.U., Nguyen, M., Desponds, J., and Mayer, A. (2020). Early life imprints the hierarchy of T cell clone sizes. arViv, arXiv:2007.11113. https://arxiv.org/abs/2007.11113.

Gillespie, D.T. (1977). Exact stochastic simulation of coupled chemical reactions. J. Phys. Chem. *81*, 2340–2361.

Goyal, S., Kim, S., Chen, I.S., and Chou, T. (2015). Mechanisms of blood homeostasis: lineage tracking and a neutral model of cell populations in rhesus macaques. BMC Biol. *13*, 85.

Guler, G.D., Tindell, C.A., Pitti, R., Wilson, C., Nichols, K., KaiWai Cheung, T., Kim, H.J., Wongchenko, M., Yan, Y., Haley, B., et al. (2017). Repression of Stress-Induced LINE-1 Expression Protects Cancer Cell Subpopulations from Lethal Drug Exposure. Cancer Cell *32*, 221–237.e13.

Gustavsson, B., Carlsson, G., Machover, D., Petrelli, N., Roth, A., Schmoll, H.J., Tveit, K.M., and Gibson, F. (2015). A review of the evolution of systemic chemotherapy in the management of colorectal cancer. Clin. Colorectal Cancer *14*, 1–10.

Hand, S.C., Denlinger, D.L., Podrabsky, J.E., and Roy, R. (2016). Mechanisms of animal diapause: recent developments from nematodes, crustaceans, insects, and fish. Am. J. Physiol. Regul. Integr. Comp. Physiol. *310*, R1193–R1211.

Hangauer, M.J., Viswanathan, V.S., Ryan, M.J., Bole, D., Eaton, J.K., Matov, A., Galeas, J., Dhruv, H.D., Berens, M.E., Schreiber, S.L., et al. (2017). Drug-tolerant persister cancer cells are vulnerable to GPX4 inhibition. Nature *551*, 247–250.

Hart, T., Chandrashekhar, M., Aregger, M., Steinhart, Z., Brown, K.R., MacLeod, G., Mis, M., Zimmermann, M., Fradet-Turcotte, A., Sun, S., et al. (2015). High-Resolution CRISPR Screens Reveal Fitness Genes and Genotype-Specific Cancer Liabilities. Cell *163*, 1515–1526.

Hata, A.N., Niederst, M.J., Archibald, H.L., Gomez-Caraballo, M., Siddiqui, F.M., Mulvey, H.E., Maruvka, Y.E., Ji, F., Bhang, H.E., Krishnamurthy Radhakrishna, V., et al. (2016). Tumor cells can follow distinct evolutionary paths to become resistant to epidermal growth factor receptor inhibition. Nat. Med. *22*, 262–269.

He, B., Zhang, H., Wang, J., Liu, M., Sun, Y., Guo, C., Lu, J., Wang, H., and Kong, S. (2019). Blastocyst activation engenders transcriptome reprogram affecting X-chromosome reactivation and inflammatory trigger of implantation. Proc. Natl. Acad. Sci. USA *116*, 16621–16630.

Hsieh, A.C., Liu, Y., Edlind, M.P., Ingolia, N.T., Janes, M.R., Sher, A., Shi, E.Y., Stumpf, C.R., Christensen, C., Bonham, M.J., et al. (2012). The translational landscape of mTOR signalling steers cancer initiation and metastasis. Nature *485*, 55–61.

Huang, W., Sherman, B.T., and Lempicki, R.A. (2009a). Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. Nucleic Acids Res. *37*, 1–13.

Huang, W., Sherman, B.T., and Lempicki, R.A. (2009b). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat. Protoc. *4*, 44–57.

Kaku, Y., Tsuchiya, A., Kanno, T., and Nishizaki, T. (2015). Irinotecan induces cell cycle arrest, but not apoptosis or necrosis, in Caco-2 and CW2 colorectal cancer cell lines. Pharmacology *95*, 154–159.

Kaldalu, N., and Tenson, T. (2019). Slow growth causes bacterial persistence. Sci. Signal. *12*, eaay1167.

Kint, C.I., Verstraeten, N., Fauvart, M., and Michiels, J. (2012). New-found fundamentals of bacterial persistence. Trends Microbiol. *20*, 577–585.

Kluin, R.J.C., Kemper, K., Kuilman, T., de Ruiter, J.R., Iyer, V., Forment, J.V., Cornelissen-Steijger, P., de Rink, I., Ter Brugge, P., Song, J.Y., et al. (2018). XenofilteR: computational deconvolution of mouse and human reads in tumor xenograft sequence data. BMC Bioinformatics *19*, 366.

Koboldt, D.C., Zhang, Q., Larson, D.E., Shen, D., McLellan, M.D., Lin, L., Miller, C.A., Mardis, E.R., Ding, L., and Wilson, R.K. (2012). VarScan 2: somatic mutation and copy number alteration discovery in cancer by exome sequencing. Genome Res. *22*, 568–576.

Korotkevich, G., Sukhov, V., and Sergushichev, A. (2019). Fast gene set enrichment analysis. bioRxiv. https://doi.org/10.1101/060012.

Kostál, V., Simůnková, P., Kobelková, A., and Shimada, K. (2009). Cell cycle arrest as a hallmark of insect diapause: changes in gene transcription during diapause induction in the drosophilid fly, Chymomyza costata. Insect Biochem. Mol. Biol. *39*, 875–883.

Kreso, A., van Galen, P., Pedley, N.M., Lima-Fernandes, E., Frelin, C., Davis, T., Cao, L., Baiazitov, R., Du, W., Sydorenko, N., et al. (2014). Self-renewal as a therapeutic target in human colorectal cancer. Nat. Med. *20*, 29–36.

Lan, X., Jörg, D.J., Cavalli, F.M.G., Richards, L.M., Nguyen, L.V., Vanner, R.J., Guilhamon, P., Lee, L., Kushida, M.M., Pellacani, D., et al. (2017). Fate mapping of human glioblastoma reveals an invariant stem cell hierarchy. Nature *549*, 227–232.

Lee, J.E., Oh, H.A., Song, H., Jun, J.H., Roh, C.R., Xie, H., Dey, S.K., and Lim, H.J. (2011). Autophagy regulates embryonic survival during delayed implantation. Endocrinology *152*, 2067–2075.

Lees, A., McIntyre, A.J., Crawford, N.T., Falcone, F., McCann, C., Holohan, C., Quinn, G.P., Roberts, J.Z., Sessler, T., Gallagher, P.F., et al. (2020). The pseudo-caspase FLIP(L) regulates cell fate following p53 activation. Proc. Natl. Acad. Sci. USA *117*, 17808–17819.

Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics *25*, 1754–1760.

Liau, B.B., Sievers, C., Donohue, L.K., Gillespie, S.M., Flavahan, W.A., Miller, T.E., Venteicher, A.S., Hebert, C.H., Carey, C.D., Rodig, S.J., et al. (2017). Adaptive Chromatin Remodeling Drives Glioblastoma Stem Cell Plasticity and Drug Tolerance. Cell Stem Cell *20*, 233–246.e7.

Liberzon, A., Birger, C., Thorvaldsdóttir, H., Ghandi, M., Mesirov, J.P., and Tamayo, P. (2015). The Molecular Signatures Database (MSigDB) hallmark gene set collection. Cell Syst. *1*, 417–425.

Mayakonda, A., Lin, D.C., Assenov, Y., Plass, C., and Koeffler, H.P. (2018). Maftools: efficient and comprehensive analysis of somatic variants in cancer. Genome Res. *28*, 1747–1756.

McCarthy, D.J., Chen, Y., and Smyth, G.K. (2012). Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. Nucleic Acids Res. *40*, 4288–4297.

Merino, D., Weber, T.S., Serrano, A., Vaillant, F., Liu, K., Pal, B., Di Stefano, L., Schreuder, J., Lin, D., Chen, Y., et al. (2019). Barcoding reveals complex clonal behavior in patient-derived xenografts of metastatic triple negative breast cancer. Nat. Commun. *10*, 766.

Min, M., and Spencer, S.L. (2019). Spontaneously slow-cycling subpopulations of human cells originate from activation of stress-response pathways. PLoS Biol. *17*, e3000178.

Mudunuri, U., Che, A., Yi, M., and Stephens, R.M. (2009). bioDBnet: the biological database network. Bioinformatics *25*, 555–556.

O'Brien, C.A., Kreso, A., Ryan, P., Hermans, K.G., Gibson, L., Wang, Y., Tsatsanis, A., Gallinger, S., and Dick, J.E. (2012). ID1 and ID3 regulate the self-renewal capacity of human colon cancer-initiating cells through p21. Cancer Cell *21*, 777–792.

Pek, M., Yatim, S.M.J.M., Chen, Y., Li, J., Gong, M., Jiang, X., Zhang, F., Zheng, J., Wu, X., and Yu, Q. (2017). Oncogenic KRAS-associated gene signature defines co-targeting of CDK4/6 and MEK as a viable therapeutic strategy in colorectal cancer. Oncogene *36*, 4975–4986.

Quayle, L.A., Ottewell, P.D., and Holen, I. (2018). Chemotherapy resistance and stemness in mitotically quiescent human breast cancer cells identified by fluorescent dye retention. Clin. Exp. Metastasis *35*, 831–846.

Rambow, F., Rogiers, A., Marin-Bejar, O., Aibar, S., Femel, J., Dewaele, M., Karras, P., Brown, D., Chang, Y.H., Debiec-Rychter, M., et al. (2018). Toward Minimal Residual Disease-Directed Therapy in Melanoma. Cell *174*, 843–855.e19.

Ramirez, M., Rajaram, S., Steininger, R.J., Osipchuk, D., Roth, M.A., Morinishi, L.S., Evans, L., Ji, W., Hsu, C.H., Thurley, K., et al. (2016). Diverse drug-resistance mechanisms can emerge from drug-tolerant cancer persister cells. Nat. Commun. *7*, 10690.

Recasens, A., and Munoz, L. (2019). Targeting Cancer Cell Dormancy. Trends Pharmacol. Sci. *40*, 128–141.

Risom, T., Langer, E.M., Chapman, M.P., Rantala, J., Fields, A.J., Boniface, C., Alvarez, M.J., Kendsersky, N.D., Pelz, C.R., Johnson-Camacho, K., et al. (2018). Differentiation-state plasticity is a targetable resistance mechanism in basal-like breast cancer. Nat. Commun. *9*, 3815.

Robinson, M.D., McCarthy, D.J., and Smyth, G.K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics *26*, 139–140.

Roesch, A., Fukunaga-Kalabis, M., Schmidt, E.C., Zabierowski, S.E., Brafford, P.A., Vultur, A., Basu, D., Gimotty, P., Vogt, T., and Herlyn, M. (2010). A temporarily distinct subpopulation of slow-cycling melanoma cells is required for continuous tumor growth. Cell *141*, 583–594.

Roulois, D., Loo Yau, H., Singhania, R., Wang, Y., Danesh, A., Shen, S.Y., Han, H., Liang, G., Jones, P.A., Pugh, T.J., et al. (2015). DNA-Demethylating Agents Target Colorectal Cancer Cells by Inducing Viral Mimicry by Endogenous Transcripts. Cell *162*, 961–973.

Scognamiglio, R., Cabezas-Wallscheid, N., Thier, M.C., Altamura, S., Reyes, A., Prendergast, A.M., Baumgärtner, D., Carnevalli, L.S., Atzberger, A., Haas, S., et al. (2016). Myc Depletion Induces a Pluripotent Dormant State Mimicking Diapause. Cell *164*, 668–680.

Seth, S., Li, C.Y., Ho, I.L., Corti, D., Loponte, S., Sapio, L., Del Poggetto, E., Yen, E.Y., Robinson, F.S., Peoples, M., et al. (2019). Pre-existing Functional Heterogeneity of Tumorigenic Compartment as the Origin of Chemoresistance in Pancreatic Tumors. Cell Rep. *26*, 1518–1532.e9.

Shakiba, N., Fahmy, A., Jayakumaran, G., McGibbon, S., David, L., Trcka, D., Elbaz, J., Puri, M.C., Nagy, A., van der Kooy, D., et al. (2019). Cell competition during reprogramming gives rise to dominant clones. Science *364*, eaan0925.

Sharma, S.V., Lee, D.Y., Li, B., Quinlan, M.P., Takahashi, F., Maheswaran, S., McDermott, U., Azizian, N., Zou, L., Fischbach, M.A., et al. (2010). A chromatin-mediated reversible drug-tolerant state in cancer cell subpopulations. Cell *141*, 69–80.

Ślusarczyk, M., Chlebicki, W., Pijanowska, J., and Radzikowski, J. (2019). The role of the refractory period in diapause length determination in a freshwater crustacean. Sci. Rep. *9*, 11905.

Sowalsky, A.G., Ye, H., Bhasin, M., Van Allen, E.M., Loda, M., Lis, R.T., Montaser-Kouhsari, L., Calagua, C., Ma, F., Russo, J.W., et al. (2018). Neoadjuvant-Intensive Androgen Deprivation Therapy Selects for Prostate Tumor Foci with Diverse Subclonal Oncogenic Alterations. Cancer Res. *78*, 4716–4730.

Spandidos, A., Wang, X., Wang, H., and Seed, B. (2010). PrimerBank: a resource of human and mouse PCR primer pairs for gene expression detection and quantification. Nucleic Acids Res. *38*, D792–D799.

Van den Bergh, B., Fauvart, M., and Michiels, J. (2017). Formation, physiology, ecology, evolution and clinical importance of bacterial persisters. FEMS Microbiol. Rev. *41*, 219–251.

Vera-Ramirez, L., Vodnala, S.K., Nini, R., Hunter, K.W., and Green, J.E. (2018). Autophagy promotes the survival of dormant breast cancer cells and metastatic tumour recurrence. Nat. Commun. *9*, 1944.

Wang, Y., Yang, L., Zhang, J., Zhou, M., Shen, L., Deng, W., Liang, L., Hu, R., Yang, W., Yao, Y., et al. (2018). Radiosensitization by irinotecan is attributed to G2/M phase arrest, followed by enhanced apoptosis, probably through the ATM/Chk/Cdc25C/Cdc2 pathway in p53-mutant colorectal cancer cells. Int. J. Oncol. *53*, 1667–1680.

West, J., and Newton, P.K. (2019). Cellular interactions constrain tumor growth. Proc. Natl. Acad. Sci. USA *116*, 1918–1923.

Wijsman, J.H., Jonker, R.R., Keijzer, R., van de Velde, C.J., Cornelisse, C.J., and van Dierendonck, J.H. (1993). A new method to detect apoptosis in paraffin sections: in situ end-labeling of fragmented DNA. J. Histochem. Cytochem. *41*, 7–12.

Wintersinger, J.A., Dobson, S.M., Dick, J.E., and Morris, Q. (2020). Pairtree: fast reconstruction of cancer evolutionary history using pairwise mutation relationships. bioRxiv. https://doi.org/10.1101/372219.

Wolf, Y., Bartok, O., Patkar, S., Eli, G.B., Cohen, S., Litchfield, K., Levy, R., Jiménez-Sánchez, A., Trabish, S., Lee, J.S., et al. (2019). UVB-Induced Tumor Heterogeneity Diminishes Immune Response in Melanoma. Cell *179*, 219–235.e21.

Yule, G.U. (1925). II.—A mathematical theory of evolution, based on the conclusions of Dr. J. C. Willis, F. R. S. Philos. Trans. R. Soc. Lond. B Biol. Sci. *213*, 21–87.

Zhang, X., Yang, H., Yue, S., He, G., Qu, S., Zhang, Z., Ma, B., Ding, R., Peng, W., Zhang, H., et al. (2017). The mTOR inhibition in concurrence with ERK1/2 activation is involved in excessive autophagy induced by glycyrrhizin in hepatocellular carcinoma. Cancer Med. *6*, 1941–1951.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Ki-67 antibody	Novus Biologicals	Cat # NB110-90592; RRID:AB_1217069
BrdU antibody	Abcam	Cat # ab6326; RRID:AB_305426
CD326 (EpCAM) antibody	Miltenyi	Cat # 130-098-118; RRID:AB_2660300
APC-conjugated mouse H-2Kb antibody	ThermoFisher	Cat # MA5-17998; RRID:AB_2539382
Phospho-Akt (Ser473) antibody	Cell Signaling Technology	Cat # 4060; RRID:AB_2315049
Akt antibody	Cell Signaling Technology	Cat # 4691; RRID:AB_915783
Phospho-4E-BP1 (Thr37/46) antibody	Cell Signaling Technology	Cat # 2855; RRID:AB_560835
4E-BP1 antibody	Cell Signaling Technology	Cat # 9452; RRID:AB_331692
Phospho-S6 (Ser240/244) antibody	Cell Signaling Technology	Cat # 5364; RRID:AB_10694233
S6 antibody	Cell Signaling Technology	Cat # 2217; RRID:AB_331355
p-ERK1/2 (T202/Y204) antibody	Cell Signaling Technology	Cat # 4370; RRID:AB_2315112
ERK1/2 antibody	Cell Signaling Technology	Cat # 9102; RRID:AB_330744
Phospho-ULK1 (Ser757) antibody	Cell Signaling Technology	Cat # 6888; RRID:AB_10829226
ULK1 antibody	Cell Signaling Technology	Cat # 8054; RRID:AB_11178668
Atg13 antibody	Cell Signaling Technology	Cat # 13273; RRID:AB_2798169
LC3B antibody	Sigma-Aldrich	Cat # L7543; RRID:AB_796155
p16 antibody	Santa Cruz	Cat # sc-1661; RRID:AB_628067
p27 antibody	BD Biosciences	Cat # 610241; RRID:AB_397636
β-actin antibody	Sigma-Aldrich	Cat # A2228; RRID:AB_476697
IRDye 680RD Donkey anti-Rabbit IgG antibody	LI-COR Biosciences	Cat # 926-68073; RRID:AB_10954442
IRDye 680RD Donkey anti-Mouse IgG antibody	LI-COR Biosciences	Cat # 926-68072; RRID:AB_10953628
IRDye 800RD Donkey anti-Rabbit IgG antibody	LI-COR Biosciences	Cat # 926-32213; RRID:AB_621848
IRDye 800CW Donkey anti-Mouse IgG antibody	LI-COR Biosciences	Cat # 926-32212; RRID:AB_621847
Biological Samples		
Patient-derived xenografts (PDX)	University Health Network BioBank	N/A
Chemicals, Peptides, and Recombinant Proteins		
Matrigel	Corning	Cat # 354262
5-Fluorouracil (5-FU)	Sigma	Cat # F6627
Leucovorin (LV; Calcium Folinate)	Santa Cruz Biotechnology	Cat # sc252837A
Oxaliplatin (OXA)	Tocris Bioscience	Cat # 2623
Irinotecan (CPT-11)	Sigma	Cat # I1406
INK 128	Selleckchem	Cat # S2811
Chloroquine	Sigma	Cat # C6628
SBI-0206965	Selleckchem	Cat # S7885
BrdU	Sigma	Cat # B5002
Ammonium chloride solution	STEMCELL	Cat # 07800
Phenol-Chloroform-Isoamyl Alcohol	Sigma	Cat # P3803
ExTaq	Takara	Cat # RR001A

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
SYTOX Blue nucleic acid stain	Invitrogen	Cat # S11348
Vybrant DiD	Molecular Probes	Cat # MP22887
Odyssey Blocking Buffer	LI-COR	Cat # 927-40003
FxCycle Violet	Invitrogen	Cat # F10347
Critical Commercial Assays		
Gentra Puregene Tissue Kit	QIAGEN	Cat # 158689
QIAamp Blood Midi Kit	QIAGEN	Cat # 51185
DNeasy Blood and Tissue Kit	QIAGEN	Cat # 69506
PicoPure RNA Isolation Kit	Applied Biosystems	Cat # KIT0214
RNeasy kit	QIAGEN	Cat # 74106
Takara SMARTer Stranded Total RNA-Seq Kit v1 - Pico Input Mammalian library kit	Takara	Cat # 635007
Annexin V-FITC apoptosis detection kit	Thermo Fisher	Cat # 88-8005-72
Deposited Data		
Raw RNaseq data	This paper	GEO: GSE145356
Whole Exome Sequencing data	This paper	EGA: EGAS00001004773
Pairtree inputs	This paper	https://github.com/morrislab/crc-dtp-ith-analysis/
Experimental Models: Cell Lines		
POP92	O'Brien et al., 2012	N/A
POP66	O'Brien et al., 2012	N/A
CSC28	This paper	N/A
POP133	This paper	N/A
HT29	ATCC	Cat# HTB-38; RRID:CVCL_0320
HEK293T	ATCC	Cat# CRL-11268; RRID:CVCL_1926
Experimental Models: Organisms/Strains		
Mouse: NOD/SCID: NOD.Cg-*Prkdc*^{scid}/J	The Jackson Laboratory	JAX:001303; RRID:IMSR_JAX:001303
Oligonucleotides		
Illumina sequencing primer: ACACTCTTTCCCT ACACGACGC TCTTCCGATCT	This paper	N/A
Barcode Custom primer: ATCGATACCGTCGA GATCCGTTCA CTAATCG	This paper	N/A
Barcode Index Primers, see Table S2	This paper	N/A
ULK1 Primer pair (see STAR Methods)	PrimerBank https://pga.mgh.harvard.edu/primerbank/	PrimerBank ID:28011640
WIPI1 Primer pair (see STAR Methods)	PrimerBank https://pga.mgh.harvard.edu/primerbank/	PrimerBank ID: 157388938c3
WIPI2 Primer pair (see STAR Methods)	PrimerBank https://pga.mgh.harvard.edu/primerbank/	PrimerBank ID: 75677338c1
WDR45 Primer pair (see STAR Methods)	PrimerBank https://pga.mgh.harvard.edu/primerbank/	PrimerBank ID: 71483644c3
OPTN Primer pair (see STAR Methods)	PrimerBank https://pga.mgh.harvard.edu/primerbank/	PrimerBank ID: 56549110c1
ATG2A Primer pair (see STAR Methods)	PrimerBank https://pga.mgh.harvard.edu/primerbank/	PrimerBank ID: 239047270c2

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
ATG13 Primer pair (see STAR Methods)	PrimerBank https://pga.mgh.harvard.edu/primerbank/	PrimerBank ID: 326806953c3
ATG16L2 Primer pair (see STAR Methods)	PrimerBank https://pga.mgh.harvard.edu/primerbank/	PrimerBank ID: 55743107c1
RPLP0 Primer pair (see STAR Methods)	Roulois et al., 2015	N/A
Recombinant DNA		
pLJM1-(barcode library)	This paper	N/A
Software and Algorithms		
Pairtree (version bd5bfc9)	Wintersinger et al., 2020	https://github.com/morrislab/pairtree
Gencode v25	The GENCODE Project	https://www.gencodegenes.org/human/release_25.html
GRCh38.p7	Genome Reference Consortium	https://www.ncbi.nlm.nih.gov/assembly/GCF_000001405.33/
HTSeq v.0.6.1	Open source	https://pypi.org/project/HTSeq/
limma v.3.38.3	Open source	https://bioconductor.org/packages/release/bioc/html/limma.html
Picard tools v.1.98	Open source	https://broadinstitute.github.io/picard/
R (v.3.5.0)	Open source	https://www.R-project.org
edgeR (v.3.24.0)	McCarthy et al., 2012; Robinson et al., 2010	https://bioconductor.org/packages/release/bioc/html/edgeR.html
STAR v.2.4.2a	Dobin et al., 2013	https://github.com/alexdobin/STAR
XenofilteR v.0.99.0	Kluin et al., 2018	https://rdrr.io/github/PeeperLab/XenofilteR/man/XenoFilteR.html
MultiQC (v1.6)	Ewels et al., 2016	https://multiqc.info/
David (v.6.7)	Huang et al., 2009b	https://david.ncifcrf.gov/
fGSEA (v.1.8.0)	Korotkevich et al., 2019	https://github.com/ctlab/fgsea
Burrows-Wheeler aligner (BWA-MEM)	Li and Durbin, 2009	http://bio-bwa.sourceforge.net/
MuTect2 (v3.6)	Benjamin et al., 2019	https://github.com/broadinstitute/gatk/releases
Varscan2 (v2.4.2)	Koboldt et al., 2012	https://sourceforge.net/projects/varscan/files/
Sequenza (v2.1.2)	Favero et al., 2015	https://cran.r-project.org/web/packages/sequenza/index.html
vcf2maf	Cerami et al., 2012	https://github.com/mskcc/vcf2maf
Maftools (v2.0.10)	Mayakonda et al., 2018	https://www.bioconductor.org/packages/release/bioc/html/maftools.html
Gillespie Algorithm	This paper	https://github.com/smcgibbo/Gillispie-Simulations
Python	Python 3	https://www.python.org/
FlowJo	FlowJo 10.0	https://www.flowjo.com/
GraphPad Prism	GraphPad 6.0	https://www.graphpad.com/scientific-software/prism/
Other		
The Cancer Genome Atlas (TCGA)	Cancer Genome Atlas, 2012	https://gdac.broadinstitute.org/
The Molecular Signatures Database (MSigDB)	Liberzon et al., 2015	https://www.gsea-msigdb.org/gsea/msigdb/collections.jsp
Acute Lymphocytic Leukemia MRD dataset	Ebinger et al., 2016	GEO: GSE83142
Prostate cancer MRD dataset	Sowalsky et al., 2018	GEO: GSE102124
Melanoma MRD dataset	Rambow et al., 2018	GEO: GSE116237

RESOURCE AVAILABILITY

Lead Contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Catherine O'Brien (cobrien@uhnresearch.ca).

Materials Availability

This study did not generate new unique reagents.

Data and Code Availability

The accession number for the RNA sequencing data reported in this paper is GEO:GSE145356. The accession number for the Whole Exome Sequencing data reported in this paper is EGA:EGAS00001004773. Code implementing the Pairtree algorithm used for these analyses is available at https://github.com/morrislab/pairtree. The Pairtree inputs used to build clone trees are available at https://github.com/morrislab/crc-dtp-ith-analysis/. The Gillespie Algorithm used for mathematical simulations performed in this study is available at https://github.com/smcgibbo/Gillispie-Simulations.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Mice

NOD/SCID mice were used in this study and were bred in our Animal Facility, originally purchased from The Jackson laboratory. Prior to all experiments, mice were allowed one week to acclimate to housing conditions in the UHN Animal Resources Centre animal facility. For maintenance of patient-derived xenografts models, cells from frozen xenograft samples were mixed (1:1) with high concentration Matrigel (Corning, Cat # 354262) and injected subcutaneously into the flanks of NOD/SCID mice (male, 6-8 weeks old). For all experiments, 5-6 weeks old female mice were used. All mice were housed in specific pathogen-free conditions and all animal experiments were reviewed and approved by the Animal Care Committee at the University Health Network, Toronto, Canada.

Human tissue and patient-derived xenografts

Human CRC tissue was obtained with informed patient consent, as approved by the Research Ethics Board at the University Health Network in Toronto, and was processed as previously described (Kreso et al., 2014). CRC models used in this study are of female origin and are listed in Table S1.

Cell lines and Primary Cell Cultures

CRC cultures established directly from patient tissue (CSC28) or from PDXs (POP66 and POP92 (O'Brien et al., 2012), and the commercially available cell line (HT29; ATCC Cat# HTB-38, RRID:CVCL_0320) were grown as spheroids in suspension culture flasks at 37°C in a humidified incubator at 5% CO_2, as previously described (Kreso et al., 2014). All cell culture models are of female origin. The culture medium contained DMEM/F-12 (GIBCO) supplemented with 2 mM GlutaMAX (GIBCO), 10 mM HEPES (GIBCO), 1 mM sodium pyruvate (HyClone), 1X non-essential amino acids (HyClone), 1% penicillin-streptomycin (HyClone), 0.5 μg/ml Fungizone (GIBCO), 1% N2 Supplement-A (STEMCELL), 0.4% NeuroCult SM1 Neuronal Supplement (STEMCELL), 4 μg/ml, Heparin (Sigma), 0.2% lipids (Sigma), 20 ng/ml EGF (Reprokine), and 10 ng/ml basic FGF (Reprokine). All models were authenticated by STR analysis and confirmed to be negative for mycoplasma.

METHOD DETAILS

DNA barcode library

Oligonucleotides comprising a 12 base pair degenerate region (the barcode) followed by two stable bases (C or G), and one of several 4 base pair library codes were synthesized with common flanking regions. Nested PCR using the common regions generated double stranded DNA which was subsequently ligated into a second-generation lentiviral vector (pLJM1) containing a puromycin resistance cassette and ZsGreen fluorescent marker. Three barcode libraries, each identifiable by a unique library code were cloned, transformed into *E. coli* and plated as a pool on solid media. At least $5x10^6$ bacterial colonies were scraped and pooled for two of the high diversity libraries. Plasmid DNA was maxi-prepped, and a sample was sequenced to confirm diversity of $> 10^6$ unique barcodes. These libraries are named Library 0 and Library 1. The third library was generated for use as standard spike-in controls. Single colonies were selected, prepped and Sanger sequenced to identify several standard barcodes.

Lentivirus production, infection, cell sorting

Lentivirus containing barcode plasmid libraries were produced in HEK293T (ATCC Cat# CRL-11268, RRID:CVCL_1926) cells. Virus was titered for both POP66 and CSC28 cultures and the desired MOI was calculated as previously described (Hart et al., 2015). CRC cultures were dissociated and filtered through a 40 μm cell strainer to remove clumps. Cells were diluted to $1.5\text{-}2 \times 10^5$ cells per ml in culture medium containing a pre-determined concentration of polybrene (3-4 μg/ml; Sigma, Cat # TR-1003) and virus was added at a

multiplicity of infection of $\leq$ 0.1 to minimize the number of cells with multiple barcodes. After a 24-hour incubation, cells were washed, then resuspended in fresh culture medium and incubated further. At ~60 hours post-infection, cells were dissociated and barcoded cells (ZsGreen+) were purified by fluorescence-activated cell sorting (FACS) using FACSAria cell sorters (BD Biosciences) and SYTOX Blue nucleic acid stain (Invitrogen) to exclude dead cells. After sorting, an aliquot of the initial barcoded cell pool was designated as time t = 0 (T0) and snap-frozen for subsequent DNA extraction. The remaining barcoded cells were seeded in a flask and expanded in culture for 8 doubling times (~256-fold representation) to provide multiple replicates with comparable starting barcode representations and the complexity to decipher pre-existing from *de novo* acquired DTP populations. Cells were passaged as necessary to avoid spheroid overgrowth. Cells were then dissociated, counted and immediately used for downstream applications or viably frozen.

Chemicals

5-FU was purchased from Sigma (Cat # F6627), LV (Calcium Folinate) from Santa Cruz Biotechnology (Cat # sc252837A), OXA from Tocris Bioscience (Cat # 2623), CPT-11 from Sigma (Cat # I1406), INK 128 from Selleckchem, (Cat # S2811), SBI-0206965 from Selleckchem (Cat # S7885) and Chloroquine from Sigma (Cat # C6628). For *in vitro* studies, CPT-11, INK 128 and SBI-0206965 were dissolved in DMSO, and Chloroquine was dissolved in water. For *in vivo* studies, 5-FU, LV, and OXA were dissolved directly in saline, whereas CPT-11 was first dissolved in DMSO (90 mg/ml) and then diluted in saline. Dilutions of stock solutions of all drugs were made in culture medium immediately prior to use *in vitro* or administration *in vivo*.

In vivo chemotherapy drug treatments

Dissociated CRC cells were mixed (1:1) with high concentration Matrigel (Corning, Cat # 354262), and injected subcutaneously ($1x10^5$ for POP66, $2x10^5$ for CSC28, $2.5x10^5$ for POP92, $1x10^6$ for POP133, $2x10^5$ for HT-29) into the hind left and right flanks of female NOD/SCID mice, 5-6 weeks of age (2 injections per mouse). When the average tumor volume reached ~100 mm^3, tumor-bearing mice were randomized into control and treatment groups based on tumor volumes, and dosing commenced on day 0. All chemotherapies were administered by intraperitoneal injection (8 weeks for POP66 and CSC28; 4 weeks for POP92, POP133 and HT-29) or until combined tumor burden reached 1.5 cm total diameter. Treatment regimens were as follows: Vehicle-Saline control (Saline only, three times per week); Vehicle-DMSO control (Saline + 5% DMSO, twice per week); 5-FU (15 mg/kg, three times per week) and LV (20 mg/kg, three times per week); OXA (1 mg/kg, twice per week); CPT-11 (45 mg/kg, twice per week); FOLFOX (5-FU + LV + OXA); FOLFIRI (5-FU + LV + CPT-11). When combination therapies were given on the same day, 5-FU and LV were administered 1 hour before OXA or CPT-11. Body weights were measured every 1-3 days over the course of treatment, and tumor growth was monitored by caliper measurements every 1-4 days until endpoint was reached and mice were euthanized.

BrdU administration

Mice bearing tumors were first treated with DMSO or CPT-11 for 2 weeks. Then BrdU (5 mg/ml; Sigma, Cat # B5002) dissolved in phosphate-buffered saline (PBS) was administered by intraperitoneal injection (50 mg/kg) three times a week for 2 additional weeks while on treatment with DMSO or CPT-11.

Tissue collection and histological analysis

The xenograft tumors harvested after sacrifice were fixed in 10% neutral buffered formalin, dehydrated, and embedded in paraffin, then serially sectioned (4 μm) thickness, and subjected to histopathological analysis. Tissue sections were stained with hematoxylin and eosin, TUNEL stain (DNA Polymerase 1 Large (Klenow) Fragment from Promega; dATP, dCTP, dGTP from Promega, Cat # M2201; Bio-11-dUTP from Cedarlane) (Wijsman et al., 1993), Ki-67 antibody (Novus Biologicals, Cat # NB110-90592, RRID:AB_1217069) and BrdU antibody (Abcam, Cat# ab6326, RRID:AB_305426) following the manufacturer's instructions. All slides were counterstained with hematoxylin and assessed under light microscopy. For IHC analysis – TUNEL expression levels were semi-quantified using immunoreactive scores (IRS range 0-12). Score definitions: 0 (score 0-3), 1+ (score 4-6), 2+ (score 7-9), 3+ (score 10-12). For Ki-67 and BrdU percent positivity, 10 fields were randomly chosen, and the average percentage of positive cells determined. All slides were assessed by a pathologist. Slides were scanned on the Aperio Scanscope AT2 Whole Slide Scanner using a 20x objective.

Digestion of xenograft tumor tissue

Tumors were excised from mice, minced with a razor blade, and incubated in medium (DMEM/F-12, GIBCO) containing collagenase A (Roche) at 37°C for 60 minutes. After enzymatic digestion, samples were diluted in wash medium and filtered through a 40 μm cell strainer to remove undigested tissue. Cells were collected by centrifugation at 450 × *g* for 6 minutes, then resuspended in ammonium chloride solution (STEMCELL, Cat # 07800) and incubated at room temperature for 6 minutes to lyse red blood cells. Cells were collected by centrifugation at 300 × *g* for 5 minutes, then resuspended in Dulbecco's PBS (DPBS) and counted. Cells were immediately used for downstream applications or viably frozen.

DNA extraction for barcode sequencing

Genomic DNA was extracted from barcoded tumors using the Gentra Puregene Tissue Kit (QIAGEN, Cat # 158689) with modifications. Briefly, 5×10^6 viably frozen tumor cells were thawed on ice and collected by centrifuging at 400 × *g* for 5 minutes. Cell pellets

and the appropriate barcode standards were resuspended and combined in Cell Lysis Solution, then Puregene Proteinase K was added (133 μg/ml final concentration) and samples were incubated at 55°C overnight. The next day, RNase A Solution was added (20 μg/ml final concentration) and samples were incubated at 37°C for 1 hour, after which Protein Precipitation Solution was added, then samples were vortexed and transferred to a 2 mL MaXtract HD tube (QIAGEN). After 1-2 rounds of extraction with Phenol-Chloroform-Isoamyl Alcohol (25:24:1; Sigma, Cat # P3803), the upper layer was transferred to a new MaXtract tube and centrifuged at 16,000 × *g* for 3 minutes. The resulting upper layer was then transferred to a new microcentrifuge tube and DNA was precipitated with isopropanol followed by a 70% ethanol wash. DNA pellets were resuspended in 100 μl nuclease-free water and incubated at 65°C for 1 hour, followed by incubation at room temperature overnight on a tube rotator. DNA was quantified using a Nanodrop spectrophotometer (Thermo Scientific). Genomic DNA was extracted from barcoded T0 cultures using the QIAamp Blood Midi Kit (QIAGEN, Cat # 51185) following the manufacturer's Spin Protocol. Frozen cell pellets containing 3-10 × 10^6 barcoded cells were used and the appropriate barcode standards were included. Eluted DNA was precipitated with ethanol in the presence of sodium chloride followed by a 70% ethanol wash. DNA pellets were resuspended in nuclease-free water and incubated at 50°C for 1 hour. DNA was quantified using a Nanodrop spectrophotometer (Thermo Scientific) and adjusted to a final concentration of 400 ng/μl.

DNA barcode amplification and sequencing

Genomic DNA for all samples is standardized to 400ng/μl in nuclease-free water. Sequencing libraries were constructed by PCR amplification using a common 3′ primer "BL Seq Amp 3′: AATGATACGGCGACCACCGAGATCT and one of 166 unique 5′ primers "BL Seq Amp 5′ XXX":

CAAGCAGAAGACGGCATACGAGAT***NNNNNN***CGATTAGTGAACGGATCTCGACGGT where the N's represent a unique sample index. Barcode index primers are provided in Table S2. Each gDNA sample was amplified in technical triplicate with unique indexes using ExTaq (Takara, Cat # RR001A) with the PCR program: 95°C for 5 minutes, 94°C for 30 s, 65°C for 30 s, 72°C for 30 s and back to step 2, 32x followed by a 5 minute hold at 72°C. PCR efficiency was assessed by running the product on a 3% agarose gel. The 137 bp barcode library band was quantified using Bio Rad Image Lab software. Equal amounts of each PCR product were pooled into batches and PAGE purified using 15% TBE PAGE gels (Novex). Purified PCR products were quantified using Qubit, pooled and run on an Illumina HiSeq2500 using version 4 chemistry and Illumina sequencing primer: ACACTCTTTCCCTACACGACGCTCTTCC GATCT and custom primer: ATCGATACCGTCGAGATCCGTTCACTAATCG for multiplexed sample ID. Samples were demultiplexed and barcode abundances analyzed.

Barcode processing and analysis

FASTQ files for each sample were analyzed using a bespoke Perl script. Each read was processed to identify one of the three expected library codes (CCAA, ACGT, or TGGA) followed by eight bases corresponding to the vector sequence (eg: ATCGATAC), allowing up to one mismatched base for each feature and ignoring quality values. Reads lacking both of these sequences were discarded. The nucleotide sequence corresponding to the barcode was then extracted as the 18 nucleotides preceding the vector sequence, and all unique barcodes were enumerated. Barcode summary files were then merged into a single matrix. Noise introduced through sequencing or PCR errors was reduced by summing counts for barcodes within a Hamming distance of two into a single barcode record, where the barcode with the highest average abundance was retained as the "parent" barcode sequence. Technical replicates were also merged by summing counts across replicates. Each sample was then normalized for sequencing depth by dividing the read count for each barcode in a given sample by the sum of the total reads. All further analysis and visualization was performed in R or Python.

Exome sequencing and analysis

Genomic DNA from viable tumor cells was extracted using the DNeasy Blood and Tissue Kit (QIAGEN, Cat # 69506) according to manufacturer's instructions. Whole exome sequencing libraries were constructed and enriched from 200 ng of genomic DNA using the Agilent SureSelect Human All Exon v7 capture kit sequenced on Illumina HiSeq2500. Sequencing reads were aligned to human genome reference (build hg38) using Burrows-Wheeler aligner (BWA-MEM) (Favero et al., 2015; Koboldt et al., 2012; Li and Durbin, 2009). Matched tumor-normal bam files were used to call somatic single nucleotide variants (SNVs) and small insertions and deletions (Indels) with MuTect2 (v3.6) and copy number profiles using Varscan2 (v2.4.2) and Sequenza (v2.1.2) (Benjamin et al., 2019; Koboldt et al., 2012; Favero et al., 2015). Mutect2 calls were annotated and converted to maf format using vcf2maf (https://github.com/mskcc/vcf2maf; Cerami et al., 2012). MAF files and copy-number profiles for each sample were then input into the inferHeterogeneity() function from the maftools (v2.0.10) R package to calculate Mutant-Allele Tumor Heterogeneity (MATH) scores (Mayakonda et al., 2018). Binomial distributions for each mutation in diploid regions of the genome were calculated using the dbinom() function in R across the full range of possible variant allele fractions (0 to 1 by 0.001 increments) as a function of depth of sequencing and alt-allele read counts. In order to test the assumptions of the t test, specifically those of normality and equal variance we performed the Shapiro-Wilk normality test and Levene's Test of Equality of Variances, respectively. Shapiro-Wilk p value = 0.58 (group 1) and p value = 0.98 (group 2) indicating that the distribution of the data is not significantly different from a normal distribution. Two-sample Levene's test p value = 0.27 indicated that the variances of these two groups are not significantly different.

Pairtree Intratumoral Heterogeneity Analysis

Pairtree code was written in Python 3. Data analysis pipeline was written in Bash. See Data S1 for details on methodology and analysis.

Purification of tumor cells from DTP-state xenografts

Viably frozen CRC cells (barcoded POP66, t = 0' sample described above) were thawed, then expanded in culture for 3 passages. Cells were then dissociated and injected as described into female NOD/SCID mice, 6 weeks of age (2 injections at 2×10^5 cells each; 5 mice for control group, 35 mice for CPT-11 group). When the average tumor volume reached 250 mm^3, tumor-bearing mice were randomized into control and treatment groups based on tumor volumes, and dosing commenced on day 0. Treatment regimens were: Vehicle-DMSO (Saline + 5% DMSO control) twice per week for 1 week, or CPT-11 (45 mg/kg) twice per week for 6-7 weeks. Body weights and tumor sizes were measured every 3-4 days until mice were euthanized and tumors were harvested. Tumors were excised, pooled for the CPT-11 group ($\geq$6 tumors per sample), and processed as described above, except that samples were digested for 30-35 minutes, then washed and filtered through a 70 μm cell strainer, and incubated in ammonium chloride solution on ice for 10 minutes.

Dissociated cells were stained and prepared for FACS as previously described with modifications. Briefly, cells were resuspended in DPBS/2% FBS to a concentration of 5×10^6 cells/ml. PE-conjugated antibodies for human CD326 (EpCAM; Miltenyi Cat # 130-098-118; RRID:AB_2660300; 1:25) and APC-conjugated antibodies for mouse H-2Kb (ThermoFisher, Cat # MA5-17998; RRID:AB_2539382; 1:20) were added, and cells with antibodies were aliquoted 100 μl/tube and incubated on ice for 60 minutes, protected from light. Cells were then collected by centrifugation at 400 × *g* for 3 minutes, washed once, then resuspended in DPBS/2% FBS with SYTOX Blue nucleic acid stain (Invitrogen, Cat # S11348), filtered, and pooled. Live human CRC cells (PE+, APC-, Sytox-) were purified by FACS using a MoFlo Astrios cell sorter (Beckman Coulter). After sorting, cells were collected by centrifugation at 450 × *g* for 15 minutes, then snap-frozen for subsequent RNA extraction.

RNA extraction

RNA was extracted from FACS-purified tumor cells was using PicoPure RNA Isolation Kit (Applied Biosystems, Cat # KIT0214) following the manufacturer's instructions. RNA was extracted from spheroid cultures was using RNeasy kit (QIAGEN, Cat # 74106) following the manufacturer's instructions. RNA was quantified using NanoDrop Spectrophotometer (Thermo Scientific). For assessment of RNA abundance in cells, cultures were treated with DMSO or CPT-11 (1 μM) for 14 days, trypsinized to single cells and counted. RNA was extracted from $5x10^5$ cells of each treatment condition and the RNA concentration (per cell) was determined.

RNA-sequencing and data processing

RNA sequencing libraries were prepared using Takara SMARTer Stranded Total RNA-Seq Kit v1 - Pico Input Mammalian library kit (Takara, Cat # 635007), loaded in equimolar amounts on an Illumina NextSeq500 sequencer, and sequenced to an average depth of 40M 75 base-pair paired-end reads per sample. Raw FASTQ files were examined for quality using MultiQC (v1.6; Ewels et al., 2016) and pre-processed to remove any reads shorter than 36bp using a bespoke Perl script. Due to potential contamination from mouse stroma, reads were first were mapped to Gencode v25 transcript models and hg38 human genome sequence using the STAR short-read aligner (v.2.4.2a; Dobin et al., 2013), and separately against mouse transcript models (Gencode vM12, mm10). Adaptor trimming was included using the following argument for STAR:–clip3pAdapterSeq AGATCGGAAGAGC. Next, the aligned BAM files generated for each sample were filtered using the R package XenofilteR (v.0.99; Kluin et al., 2018) to flag and remove reads of mouse origin. Finally, the filtered BAM files were summarized to gene-level read counts using HTSeq (v.0.6.1) with the following options: -f bam -r pos -s reverse -t gene < BAM file > . All read count files were merged together into a matrix for further analysis.

RNA-seq bioinformatic analysis

Count data were subsequently imported into R (v.3.5.0) and filtered to remove non-expressed genes. For read-depth normalization, filtered gene count data were analyzed with the limma using the edgeR package v.3.24.0 in R (see Tables S3 and S4; McCarthy et al., 2012; Robinson et al., 2010). Thresholds for differential expression were fold-change > 1.5 and FDR < 0.05. Gene ontology analysis was performed with David 6.7 (https://david-d.ncifcrf.gov/) (Huang et al., 2009a, 2009b). The full list of GO terms (FDR < 0.05) is provided in Table S5. Other RNA-seq analyses and statistics were performed in R utilizing custom R scripts. Control samples noted in Figure 4 are DMSO and Saline treated tumors.

RNA-seq data from ESCs (control and paused) (Bulut-Karslioglu et al., 2016) and embryos (control E4.5 and diapause) (Boroviak et al., 2015) were also filtered to remove non-expressed genes, read-depth normalized and analyzed with limma using the ''edgeR'' package in R. For direct comparison with tumor expression data, normalized gene expression counts were first transformed into z-scores separately for each dataset before merging the data. Genes with matching nomenclature were considered as orthologs (see Table S4). Principal Component Analysis (PCA) was performed with the prcomp function using the z-scores (normalized within each dataset) and plotted with the function xyplot in R. Hierachical clustering of tumors and embryos was performed with the ''ward.D'' method of the hclust function in R using the z-scores (normalized within each dataset).

For gene set enrichment analysis (GSEA), gene signatures were obtained from the MSigDB "Hallmark gene sets" collection (Liberzon et al., 2015). All expressed genes were pre-ranked by t value (DTP state/control, or resistant tumors/control, performed with edgeR as previously described). GSEA analysis was then performed using the fGSEA package (Korotkevich et al., 2019). The full list of Hallmark pathways is provided in Table S5. For comparison purposes, a literature search was also performed on NCBI Gene Expression Omnibus (GEO, https://www.ncbi.nlm.nih.gov/geo/) to acquire expression data from models of low proliferation, with a focus on cell cycle arrest, and the following models were selected: ESCs grown in MEK/GSK inhibitor-containing "2i" media, Nutlin3a-treated HCT116 colon cancer cells and neural crest cells, palbociclib (CDK4/6 inhibitor)-treated DLD1 colon cancer cells, spontaneous p21-arrested human mammary epithelial cell line MCF10A, and Rb overexpression-arrested ESCs. GEO accession numbers: GSE81285, GSE113682, GSE128191, GSE74603, GSE122927 and GSE29783 (Bulut-Karslioglu et al., 2016; Lees et al., 2020; Bowen et al., 2019; Pek et al., 2017; Min and Spencer, 2019; Conklin et al., 2012). For each model of cell cycle arrest, as well as the embryonic pausing models (paused ESCs and diapaused embryos), differential analysis and GSEA were performed as previously described for the DTP state. Then, the normalized enrichment scores (NES) for all models were merged. In order to focus on relevant pathways and avoid issues linked to the non-continuous distribution of the NES, scores with FDR > 0.05 were ignored and replaced by missing values (NAs). The data were then plotted as a heatmap, using the "ward.D" method of the hclust function for clustering purposes, with missing values colored in gray.

For direct comparison of the DTP state and paused ESCs by GSEA, the top differentially expressed genes in paused ESCs ($t > 10$ or $t < -10$, as computed with edgeR) were selected as gene lists. Then, all expressed genes were pre-ranked by t value for the DTP state and GSEA analysis was performed using the fGSEA package (Korotkevich et al., 2019).

To establish a general "embryonic pausing signature," differential expression was first measured in each embryonic model independently (i.e., paused ESC/control ESC and diapaused embryos/E4.5 embryos). Genes significantly dysregulated in both models (upregulated in both or downregulated in both, with FDR < 0.05 for each model, as computed with edgeR) were selected for the signature (see Table S6). Then, the expression value (z-score) was multiplied by −1 if the gene was downregulated in the embryonic models, while the value of genes upregulated in the embryonic models was kept unchanged, so that a positive expression value would always reflect a change similar to embryonic pausing. For each sample, the signature score was then computed as the mean of all genes. The same method was applied to datasets related to MRD, obtained from the NCBI GEO database (GSE83142, GSE102124 and GSE116237) (Ebinger et al., 2016; Rambow et al., 2018; Sowalsky et al., 2018). When needed, gene ID conversion was performed using bioDBnet (https://biodbnet-abcc.ncifcrf.gov/db/db2db.php; Mudunuri et al., 2009).

Gene signatures for the "mTORC1 response" and the "Myc response" were obtained from the MSigDB "Hallmark gene sets" collection (Liberzon et al., 2015). For each sample, the score was computed as the average expression z-score of all signature genes. The list of "diapause-related autophagy genes" was obtained from He 2019 (He et al., 2019).

Survival analysis

For the survival analysis, we utilized The Cancer Genome Atlas (TCGA) CRC cohort (Cancer Genome Atlas, 2012). Gene expression (RSEM z-scores) and overall survival information was retrieved from the "Firehose" portal (https://gdac.broadinstitute.org/) for colon adenocarcinoma (COAD) and rectal adenocarcinoma (READ), and merged (COREAD). Embryonic pausing signature score was computed for each sample by using differentially expressed genes. As described in the previous section, first the expression value of each gene (z-score) was multiplied by its sign from the differential gene expression analysis (−1 for downregulated, +1 for upregulated genes). Next, the embryonic pausing signature was computed by averaging the values. To determine the cut-off between high and low expression, all percentiles were computed, and the best-performing threshold (in terms of the p value for the corresponding observation, regardless of the direction of the change) between the two groups was selected as the cut-off. We then applied the Kaplan-Meier model and log rank test was used to determine the p value.

For autophagy, the gene signature was obtained from "GO_REGULATION_OF_AUTOPHAGY" from the Gene Ontology Consortium (GO:0010506). The autophagy signature score was computed with the same method as previously described for the pausing signature, with the exception that z-scores were not multiplied by +1/-1. Kaplan-Meier model and log rank test were then used to evaluate the effects of autophagy on survival in the CRC cohort (COREAD). We applied the same method to other TCGA cohorts to identify cancer types in which autophagy was associated with poor survival.

In vitro chemotherapy drug treatments

CRC cells were treated *in vitro* with the chemotherapies at these indicated doses: CPT-11 (1 μM), INK 128 (25 nM), SBI-0206965 (2.5 μM) and Chloroquine (20 μM).

Cell growth analysis

For cell growth analysis by counting, cells were plated in multiple T25 flasks ($0.5x10^6$ cells for POP66 and CSC28; $1x10^6$ for POP92 and HT29) in 7 mL of medium and treated with DMSO (vehicle control), CPT-11 (1 μM), INK 128 (25 nM), SBI-0206965 (2.5 μM) or combination (CPT-11+SBI-0206965) at the indicated treatment concentrations. Cells were treated twice a week for the indicated durations. To count, cells from 1 flask of each indicated treatment were dissociated and counted manually using a hemocytometer after trypan blue exclusion. All cells under treatment were subjected to brief trypsinization every 7-10 days to maintain spheroid growth in logarithmic phase.

Fluorescent dye label retention assay

CRC cultures were labeled with Vybrant DiD (Molecular Probes, Cat # MP22887) according to manufacturer's instructions. Labeled samples were treated with either DMSO or CPT-11 (1 μM) for 14 days (2 times per week). At days 0, 7 and 14 on treatment, and day 24 (10 days post-treatment cessation), cultures were dissociated, labeled with SYTOX Blue nucleic acid stain (Invitrogen, Cat # S11348) and were assessed for the degree of Vybrant DiD staining present within each sample. Cytofluorimetric analysis of treated cultures post-DiD staining was done on the BD LSR-II (BD Biosciences). Cells were initially gated for live cell population (sytox negative stained cells) and then for Vybrant DiD-negative (DiD-) and DiD-positive (DiD+) cells. Negative control samples were cells that were not labeled with Vybrant DiD and were used to set the gate to detect DiD− events. Subsequently, post-labeling treated cells that were in this DiD− gate were determined to have entirely lost their initial label. Cells not found in this gate were determined to be positive for Vybrant DiD dye label. Percentage of cells found in each gate was determined at the indicated time points using FlowJo software (version 10) and statistical analysis was performed using the GraphPad Prism software (version 6).

Real-time PCR

Quantitative real-time PCR was performed as previously described using Superscript III (Invitrogen, Cat # 18080093) to amplify cDNA and SensiFAST SYBR Hi-ROX qPCR kit (FroggaBio, Cat # BIO-92020) for real-time PCR (O'Brien et al., 2012). Housekeeping gene RPLP0 was used for normalization. Primer sequences: ULK1 Fwd: GGACACCATCAGGCTCTTCC, Rev: GAAGCCGAAGT CAGCGATCT (PrimerBank ID:28011640); WIPI1 Fwd:AACAGGTCTATGTGCTCTCTCT, Rev: CTCATGGGCAGCAATAGTGC (PrimerBank ID: 157388938c3); WIPI2 Fwd: CCATCGTCAGCCTTAAAGCAC, Rev: TCCAGGCATACTATCAGCCTC (PrimerBank ID: 75677338c1); WDR45 Fwd: GAGAAGCAACTGCTAGTGTTCC, Rev: GGCTGGTTTAGAGACACACAG (PrimerBank ID: 71483644c3); OPTN Fwd: CCAAACCTGGACACGTTTACC, Rev: CCTCAAATCTCCCTTTCATGGC (PrimerBank ID: 56549110c1); ATG2A Fwd: TGTCCCTGTAGCCATGTTCG, Rev: TCAGGATCTCCGTGTACTCAG (PrimerBank ID: 239047270c2); ATG13 Fwd: TCACTTTGTGGACCGTCCCTA, Rev: TGGTACACACTTCTTGAGAGTCT (PrimerBank ID: 326806953c3); ATG16L2 Fwd: TGGA CAAGTTCTCAAAGAAGCTG, Rev: CCTCAGTGCGACCAGTGAT (PrimerBank ID: 55743107c1); RPLP0 Fwd: CAGACAGACACTG GCAACA, Rev: ACATCTCCCCCTTCTCCTT (Spandidos et al., 2010; Roulois et al., 2015).

Western blot analysis

Cultures were pelleted, washed with ice-cold PBS and then lysed in RIPA buffer (Thermo Fisher, Cat # 89900) supplemented with protease/phosphatase inhibitors (Cell Signaling, Cat # 5872) for 15 minutes at 4°C. Samples were then sonicated in a Q800 Water-bath sonicator (QSonica) for 15 cycles of 8 s ON, 15 s OFF. Lysates were cleared after centrifugation at 16000 x *g* for 15 min, protein concentration was determined using Pierce BCA Protein Assay kit (Thermo Fisher Scientific, Cat # 23225) and lysates were then boiled in Laemmli buffer for 5 minutes (Bio-Rad, Cat # 1610747). Proteins were resolved by SDS-PAGE (Bio-Rad, Cat # 456-1094) and transferred onto polyvinylidene fluoride membranes (Millipore, Cat # IPFL00010). Membranes were blocked for 1 hour in Odyssey Blocking Buffer (LI-COR, Cat # 927-40003), and subjected to immunobloting overnight at 4°C with the following primary antibodies: Phospho-Akt (Ser473, Cell Signaling Tech., Cat #4060, RRID:AB_2315049, 1:1000), Akt (Cell Signaling Tech., Cat #4691, RRID:AB_915783, 1:1000), Phospho-4E-BP1 (Thr37/46, Cell Signaling Tech., Cat #2855, RRID:AB_560835, 1:1000), 4E-BP1 (Cell Signaling Tech., Cat #9452, RRID:AB_331692, 1:1000), Phospho-S6 (Ser240/244, Cell Signaling Tech., Cat #5364, RRID:AB_10694233, 1:1000), S6 (Cell Signaling Tech., Cat #2217, RRID:AB_331355, 1:1000), p-ERK1/2 (T202/Y204, Cell Signaling Tech., Cat #4370, RRID:AB_2315112, 1:1000), ERK1/2 (Cell Signaling Tech., Cat #9102, RRID:AB_330744, 1:1000), Phospho-ULK1 (Ser757, Cell Signaling Tech., Cat #6888, RRID:AB_10829226, 1:1000), ULK1 (Cell Signaling Tech., Cat #8054, RRID:AB_11178668, 1:1000), Atg13 (Cell Signaling Tech., Cat #13273, RRID:AB_2798169, 1:1000), LC3B (Sigma, Cat #L7543, RRID:AB_796155, 1:1000), p16 (Santa Cruz, Cat # sc-1661, RRID:AB_628067, 1:250), p27 (BD Biosciences, Cat # 610241, RRID:AB_397636, 1:500), β-actin (Sigma, Cat #A2228, RRID:AB_476697, 1:2000). Bound antibodies were visualized using IRDye secondary antibodies (LI-COR, Cat # 926-68073, 926-32213, 926-68072, 926-32212; RRID:AB_10954442, RRIDs:AB_621848, AB_10953628, AB_621847) on an Odyssey Classic Infrared Imaging System (LI-COR).

Cell cycle and apoptosis/necrosis

For cell cycle analysis, cells were treated for 14 days with DMSO, CPT-11 (1 μM), trypsinized, washed and fixed in ice-cold ethanol overnight. Cells were then washed with PBS twice and resuspended in FxCycle Violet (Invitrogen, Cat # F10347) following the manufacturer's instructions. For apoptosis/necrosis analysis, Annexin V-FITC and propidium iodide (PI) staining (Annexin V-FITC apoptosis detection kit, Thermo Fisher, Cat # 88-8005-72) was performed following the manufacturer's instructions. For both assays, cells were analyzed on BD LSR-II (BD Biosciences). Data was analyzed with FlowJo v10.

Mathematical modeling, stochastic simulations

Here we describe the quantitative analysis of clonal response of lentiviral barcoded human CRC xenografts to chemotherapy treatment. Our goal is to infer rules of cancer cell dynamics leading to the observed heterogeneity in clonal output of individual cells. The central question we address is: Does clonal heterogeneity come from (a) ''selective dynamics'' from fitness difference among clones, or (b) from the stochastic dynamics of an ''equipotent'' population of tumorogeneic cells? We find that the experimental data is consistent with an equipotent scenario, and the shape of the barcode frequency distribution can be explained by a model

with proliferative hierarchy (Goyal et al., 2015; Lan et al., 2017) and sub-exponential (power law) tumor growth kinetics (West and Newton, 2019).

Nature of barcode frequency distribution across treatments

To get a quantitative understanding of clonal dynamics during tumor growth we analyzed the barcode frequency distribution obtained from the experiments. Here, we show that the barcode frequency distribution exhibits a log-linear relationship, i.e., a linear relationship on a log-log plot: $p(n) \sim n^{-(1+\alpha)}$. We estimated the value of α using maximum likelihood estimation (Clauset et al., 2009). In agreement with the maximum likelihood estimates of α, the corresponding cumulative distribution $Q(n) = \sum_{n'>n} p(n') \sim n^{-\alpha}$ also exhibits a log-linear relationship (see Figures 3 and S5).

Origin of log-linear distribution in the context of population dynamics from interactions between clones

Cellular interactions are thought to constrain tumor growth, causing an initially exponentially growing tumor to slow down and exhibit power law growth kinetics (West and Newton, 2019). If cancer growth kinetics follow power law kinetics, i.e., $\dot{n} = n^{1-\alpha} + \theta C$, where α quantifies the interaction between dividing cells, θ is the rate for a particular clone to transition to rapidly cycling state, and C is total number of progenitor like engrafted cells each with a unique barcode, then the population size at long times grows polynomially, $n(t) \sim (\alpha t)^{1/\alpha}$. Now, for the i^{th} clone, its size c_i follows mean-field kinetics given by, $\dot{c} = c/n^{\alpha} = c/\alpha t$, giving $c(T-t) = (T/t)^{1/\alpha}$. Assuming independent and stochastic entry of each clone, the probability distribution for a clone to get to a particular size is, $p(n, T) \approx \int_0^T dt \theta e^{-\theta t} \delta(n - c(T-t))$. Although, this integral ignores the stochasticity in the birth-death process of the fast-cycling cells, it still highlights the source of the log-linear relationship. The integral is readily evaluated via a change of variables $y = c(T-t)$ leading to $t = Ty^{-\alpha}$, giving $p(n,T) \sim \int dy y^{-(1+\alpha)} \delta(n-y) \sim n^{-(1+\alpha)}$ for large clones.

Origin of log-linear distribution in the context of population dynamics from selective dynamics

To understand how selective dynamics leads to log-linear distribution, let us consider clones with heterogeneous rates of proliferation. For a clone i undergoing a stochastic birth-death process its clone size distribution is given by $p_i(n) = \lambda_i e^{-n\lambda_i}$ where λ_i depends on its proliferative capacity and time. If we assume a generic Gamma distribution for the heterogeneity for λ, i.e., $p(\lambda) = a^b e^{-b\lambda} \lambda^{a-1}/\Gamma(a)$, the overall clone size distribution is given by $p(n) = \int d\lambda \lambda e^{-n\lambda} p(\lambda) = \frac{a/b}{(n/b+1)^{1+a}}$, which asymptotically behaves in a log-linear fashion, $p(n) \sim n^{-(1+a)}$.

Detailed Mathematical Analysis

Here we present derivation of different models that provide the basis for the heuristic analysis presented above.

(i) No proliferative hierarchy – Let r be the birth-rate of an injected cell and μ be its death rate. The outcome of such a process is summarized into a distribution of tumor having n cells at time t starting from one cell.

$$p(n,t|1,0) = \left(1 - p_S(t)\right)\delta_{n,0} + p_S(t)\frac{(1-a)}{a}a^n(1-\delta_{n,0}) \tag{1}$$

where $a = (A_1 - 1)/(A_1 - \nu_0)$ and $p_S = A_1(1-\nu_0)/(A_1 - \nu_0)$ where $A_1 = e^{(r-\mu)t}$ and $\nu_0 = \mu/r$.

(ii) Proliferative hierarchy: If there is a proliferative hierarchy, to calculate the distribution of having n number of cells starting with one injected cell, $P(n,T|1)$ is given by a combination of three mutually exclusive processes: (i) the tumor dies that is $n = 0$ if the first event for the initial starting cell is death, i.e., $\int_0^T e^{-(\theta+\mu)t}\mu dt$, (ii) nothing happens, and (iii) the first event is differentiation after which growth of progenitor cell controls the fate of the tumor, $P_+(n,T|1)$; these three contributions are mutually exclusive and are separated in Equation 2 for clarity. Here θ is the transition rate between the states.

$$P(n,T|1) = \underbrace{(1-p_S)\left(1-e^{-(\theta+\mu)T}\right)\delta_{n,0}}_{I} + \underbrace{e^{-(\alpha+\mu)T}\delta_{n,1}}_{II} + \underbrace{\int_0^T e^{-(\theta+\mu)t} \times (\theta dt) \times \left(e^{-r(T-t)}\left(1-e^{-r(T-t)}\right)^{n-1}\right)}_{III} \tag{2}$$

The non-trivial term when the first event is differentiation can be calculated readily:

$$P_+(n,T|1) = \theta \int_0^T dt e^{-(\theta+\mu)t} e^{-r(T-t)}\left(1-e^{-r(T-t)}\right)^{n-1} \tag{3}$$

$$= \frac{\theta}{r} e^{-(\theta+\mu)T} \int_0^{1-A^{-1}} dy \frac{y^{n-1}}{(1-y)^{\nu}} \tag{4}$$

$$= \frac{\theta}{r} e^{-(\theta+\mu)T} B(A;\nu,n-1) \tag{5}$$

where $\nu \equiv (\theta+\mu)/r$ and $A = e^{rT}$.

Let us evaluate $B(A;\nu,n)$ for large A and $n>0$ and $\nu \neq 1$ using integration by parts.

$$B(A;\nu,n) = \int_0^{1-1/A} dy y^n (1-y)^{-\nu} \tag{6}$$

$$= -y^n \frac{(1-y)^{-\nu+1}}{1-\nu}\Big|_0^{1-1/A} + \frac{n}{1-\nu}\int dy y^{n-1}(1-y)^{-\nu+1} \tag{7}$$

$$= -(1-1/A)^n \frac{A^{\nu-1}}{1-\nu} + \frac{n}{1-\nu} B(A;\nu-1,\, n-1) \tag{8}$$

Recursion gives,

$$B(A;\nu,n) = -\sum_{k=0}^{n} (1-1/A)^{n-k} A^{\nu-k-1} \frac{\Gamma(n+1)}{\Gamma(n-k+1)} \frac{\Gamma(1-\nu)}{\Gamma(k+2-\nu)} + \Gamma(n+1)\frac{\Gamma(1-\nu)}{\Gamma(n+1-\nu)} B(A;\nu-n,0) \tag{9}$$

Finally giving,

$$B(A;\nu,n) = -\sum_{k=0}^{n} (1-1/A)^{n-k} A^{\nu-k-1} \frac{\Gamma(n+1)}{\Gamma(n-k+1)} \frac{\Gamma(1-\nu)}{\Gamma(k+2-\nu)} + \Gamma(n+1)\frac{\Gamma(1-\nu)}{\Gamma(n+2-\nu)} \left(1 - A^{\nu-n-1}\right) \tag{10}$$

An asymptotic approximation for $n \gg 1$ and $A \to \infty$ is given by

$$P(n,T|1) \approx (1-p_S)\left(1-e^{-(\alpha+\mu)T}\right)\delta_{n,0} + e^{-(\alpha+\mu)T}\delta_{n,1} + e^{-(\alpha+\mu)T} p_S \nu \frac{1}{n^{1-\nu}} (1-\delta_{n,0}) \tag{11}$$

Note that the probability distribution for large n behaves as $n^{-(1-\nu)}$, which is what we observe for the T0 distribution, indicated by the negative slope in Figures 3C and 3D.

Crowding – Power Law Growth Kinetics

If we assume crowding reduces the growth in a population dependent way, then we can assume the growth rate to be r/n^{α}, where $n(t)$ satisfies equation $\dot{n} = r n^{1-\alpha}$, and solves to $n(t) = (rt\alpha+1)^{1/\alpha}$. Now for a birth process with r/n^{α} as the time dependent birth rate, $p(n,T|1,t) = e^{-R(t,T)}(1-e^{-R(t,T)})^{n-1}$, where $R(t,T) = \int_t^T r/n(t')^{\alpha} dt' = \log((rT\alpha+1)/(rt\alpha+1))^{\frac{1}{\alpha}}$.

Putting this together for *proliferative hierarchy with crowding* model, we get for $n \gg 1$

$$P(n,T|1,0) = \int_0^T dt \theta e^{-(\theta+\mu)t} e^{-R(t,T)} \left(1-e^{-R(t,T)}\right)^{n-1} \tag{12}$$

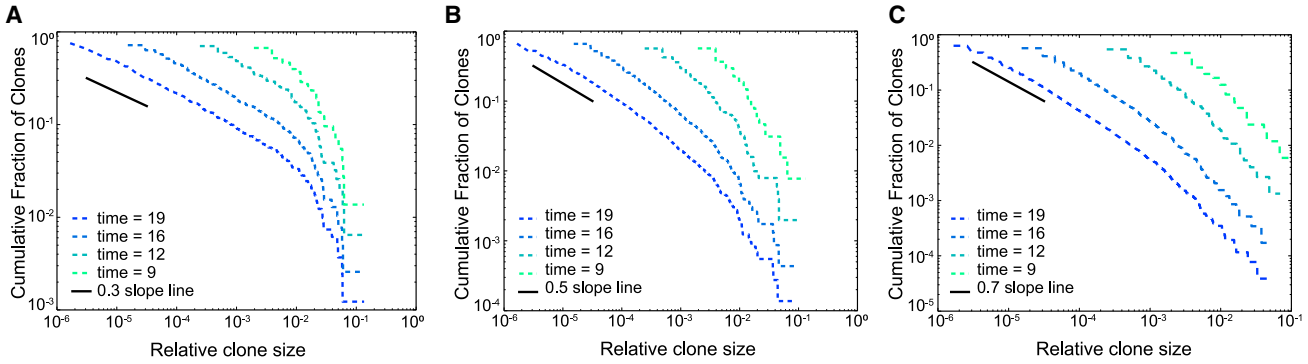
$$\sim \int_0^{1-1/A} dy e^{-(\theta+\mu)T(1-y)^{\alpha}} \frac{y^{n-1}}{(1-y)^{-\alpha}} \tag{13}$$

$$\sim B(A;-\alpha,n-1) \tag{14}$$

which behaves as $n^{-(1+\alpha)}$. Hence, the cumulative distribution behaves as $n^{-\alpha}$.

Stochastic simulation

We performed stochastic simulations using Gillespie algorithm (Gillespie, 1977) to test the validity of our mathematical analysis (https://github.com/smcgibbo/Gillispie-Simulations). With the aim to understand the role of power law kinetics on clone size distribution, we assume mean size of tumor to follow $n(T) \sim T^{1/\alpha}$, and assume growth rate of a clone to be proportional to $1/n(t)^{\alpha}$, such that the mean-field dynamics of each clone follows $\dot{c}_i \sim c_i/n(t)^{\alpha}$. The cumulative distributions obtained from the simulation (model figure, below), agrees quantitatively with the mathematical analysis and shows a log-linear relationship for clone size distribution. Here we have focused on early time population dynamics where a tumor's net growth is dominated by power law kinetics. At longer times, as tumor progresses, the steady state distribution of clones will depend on the tumor's long-term birth and death kinetics. The model figure depicts time evolution of cumulative distributions obtained from simulation of 10^5 clones with growth kinetics characterized by power law with (A) $\alpha = 0.3$, (B) $\alpha = 0.5$, and (C) $\alpha = 0.7$. Time is measured in number of generations and the solid line shows the slope of α on the log-log plot.



Model Figure: Cumulative Distributions of Simulated Power Law Growth Kinetics

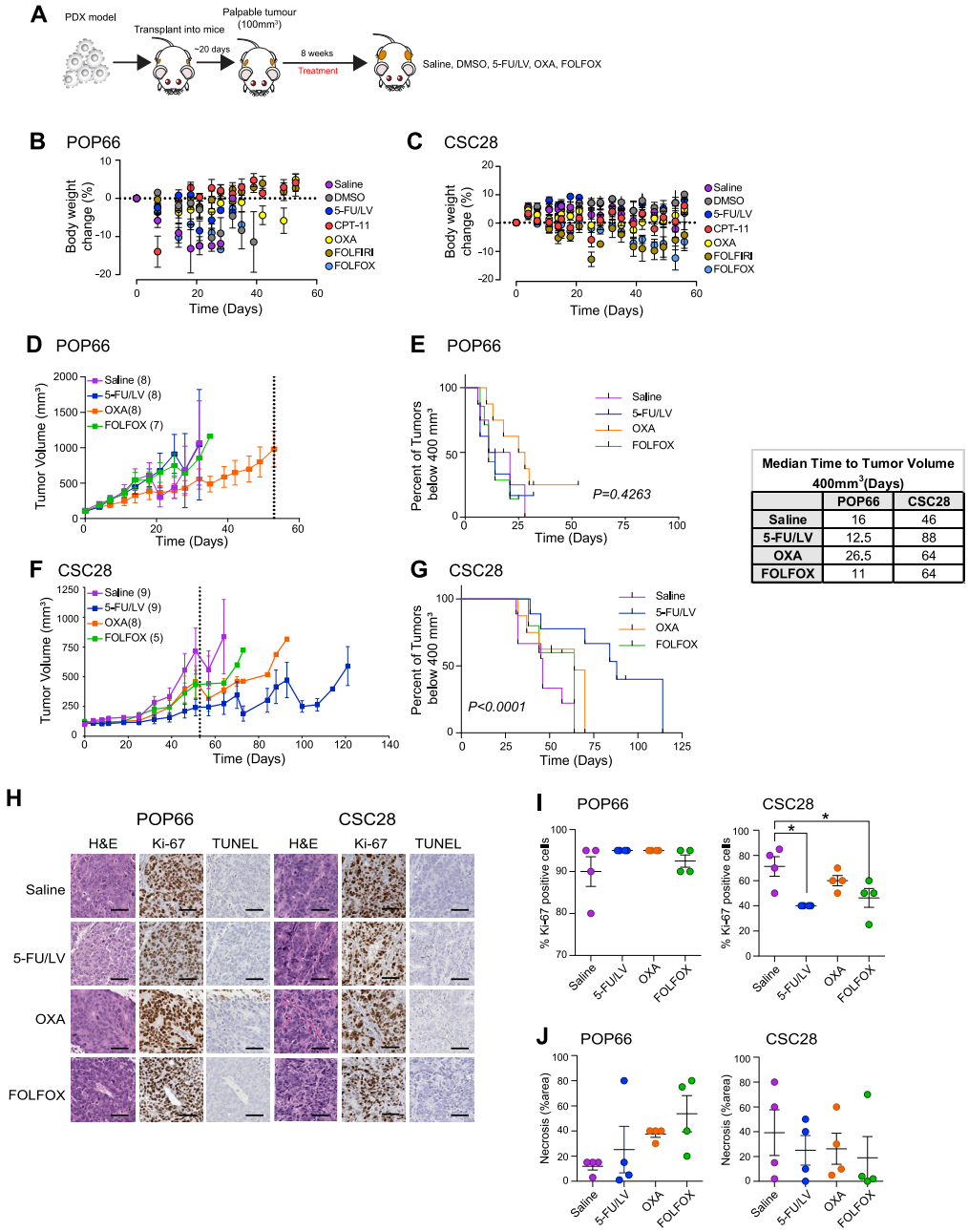
QUANTIFICATION AND STATISTICAL ANALYSIS

All the statistical details of experiments can be found in the figure legends, including the significance computed, statistical test used, and number of independent experiments performed for each assay (n). The results of quantitative data are presented as the mean ± standard deviation (SD) or standard error of mean (SEM), as indicated.

Statistical analyses were performed as indicated with each assay in the Method Details section. Statistical differences between data pairs were analyzed by Student's t test. For multiple comparisons, one-way or two-way ANOVA was used, as indicated. For barcode composition analysis, Wilcoxon rank-sum test with Holm multiple testing correction was used. For Kaplan-Meier analysis, the log rank test was used. Differences were considered statistically significant when $p < 0.05$. Significance thresholds for differential gene expression or pathway analyses were FDR < 0.05.

For MATH score analysis, to test the assumptions of the t test, specifically those of normality and equal variance, we performed the Shapiro-Wilk normality test and Levene's Test of Equality of Variances, respectively. Shapiro-Wilk p value = 0.58 (group 1) and p value = 0.98 (group 2) indicating that the distribution of the data is not significantly different from a normal distribution. Two-sample Levene's test p value = 0.27 indicated that the variances of these two groups are not significantly different.

Supplemental Figures



Median Time to Tumor Volume 400mm³(Days)

	POP66	CSC28
Saline	16	46
5-FU/LV	12.5	88
OXA	26.5	64
FOLFOX	11	64

Figure S1. Response of CRC PDXs to Chemotherapy Treatment, Related to Figure 1 and Table S1

A: Schematic of *in vivo* treatment for two CRC patient-derived xenograft (PDX) models (POP66 and CSC28). B and C: Body weight of NOD/SCID mice bearing POP66 CRC PDXs (B) from Figure 1B and (D), and CSC28 CRC PDXs (C) from Figure 1D and (F) over the course of the treatment. Data are body weight change calculated as the average percentage of the body weight of mice in each treatment group at day 0 of treatment ± SEM. D and E: Tumor growth curves (D) and Kaplan-Meier survival curves (E) of POP66 PDXs treated with Saline (vehicle control), or chemotherapy 5-FU/LV, Oxaliplatin (OXA), or the combination (FOLFOX). Dotted line indicates when treatment was stopped. Each point on growth curve represents the mean tumor volume ± SEM. F and G: Tumor growth curves (F) and Kaplan-Meier survival curves (G) of CSC28 PDXs treated with Saline (vehicle control), 5-FU/LV, OXA, or FOLFOX. Each point on growth curve represents the mean tumor volume ± SEM. Numbers in parentheses indicate biological replicates in that group. Dotted line indicates when treatment was stopped. Median time (days) to reach tumor volume 400mm^3 is listed for these POP66 and CSC28 PDXs. Statistical significance determined by log rank test (E and G). H: IHC (H&E, Ki-67, TUNEL stain, 20x) analysis of tumors from (D) and (F). Scale bar is 50 μm. I and J: Percent Ki-67 positive cells (I) and percent area Necrosis (J) are plotted. Statistical significance determined by one-way ANOVA. *p < 0.05.

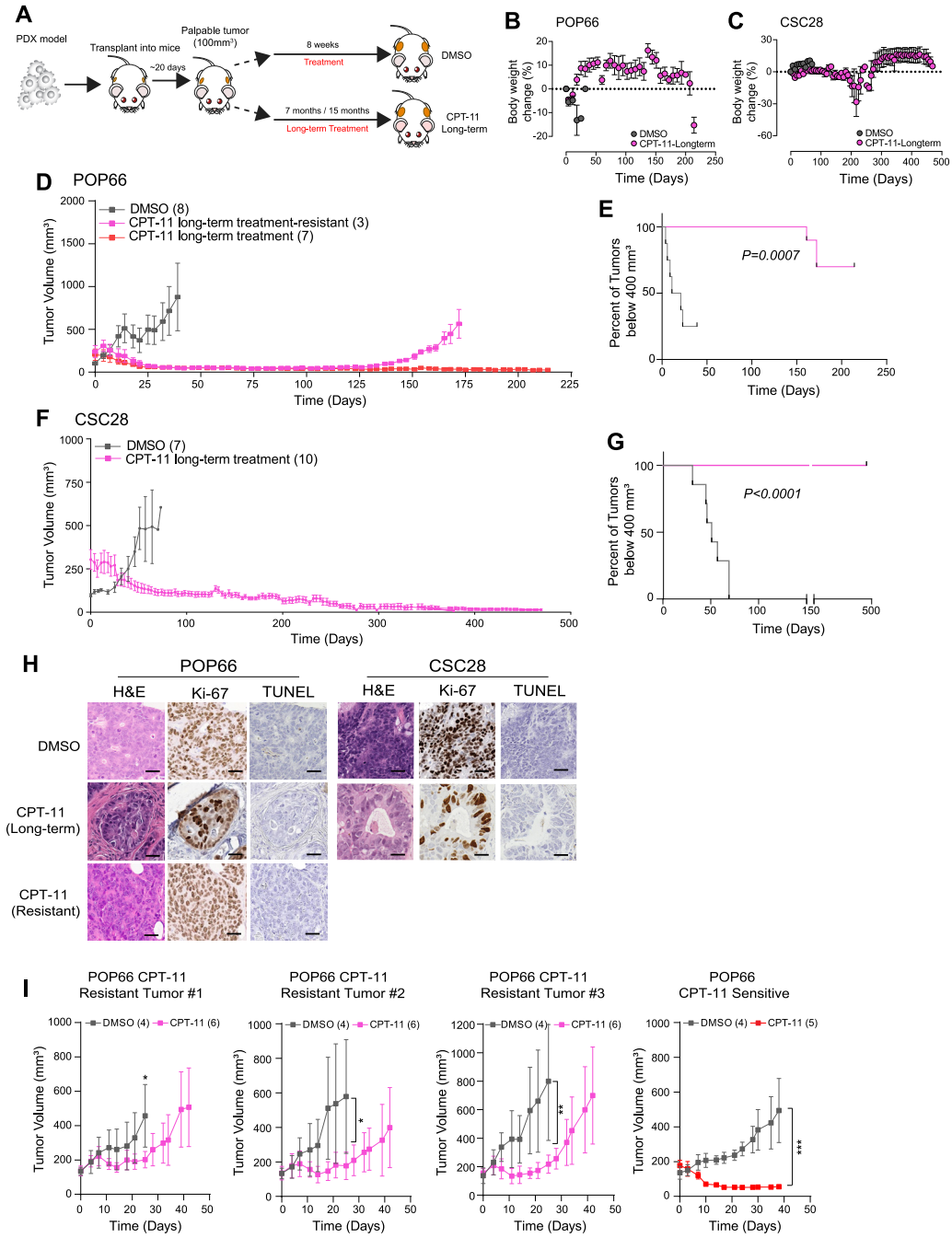


Figure S2. Long-Term Chemotherapy Treatment of CRC PDXs, Related to Figure 1

A: Schematic of *in vivo* long-term treatment for POP66 and CSC28 CRC PDX models. B and C: Body weight of NOD/SCID mice bearing POP66 (B) and CSC28 (C) CRC PDXs over the course of the treatment. Data are body weight change calculated as the average percentage of the body weight of mice in each treatment group at day 0 of treatment ± SEM. D and E: Tumor growth curves (D) and Kaplan-Meier survival curves (E) of POP66 PDXs treated with DMSO (vehicle control) or treated long-term with CPT-11 (214 days). Each point on growth curve represents mean tumor volume ± SEM. Statistical significance determined by log rank test. F and G: Tumor growth curves (F) and Kaplan-Meier survival curves (G) of CSC28 PDXs treated with DMSO (vehicle control) or treated long-term with CPT-11 (469 days). Each point on growth curve represents mean tumor volume ± SEM. Numbers in parentheses indicate biological replicates in that group. Statistical significance determined by log rank test. H: Representative IHC (H&E, Ki-67, TUNEL stain, 20x) images of DMSO, CPT-11 Long-term treated and CPT-11 Resistant tumors. Scale bar is 50 μm. I: Reinjection of the three POP66 CPT-11-Resistant tumors that grew on treatment (from D) and re-treatment with DMSO (vehicle control) or CPT-11. Each point on growth curve represents mean tumor volume ± SEM, t test. Numbers in parentheses indicate biological replicates in that group. *p < 0.05, **p < 0.01, ***p < 0.001.

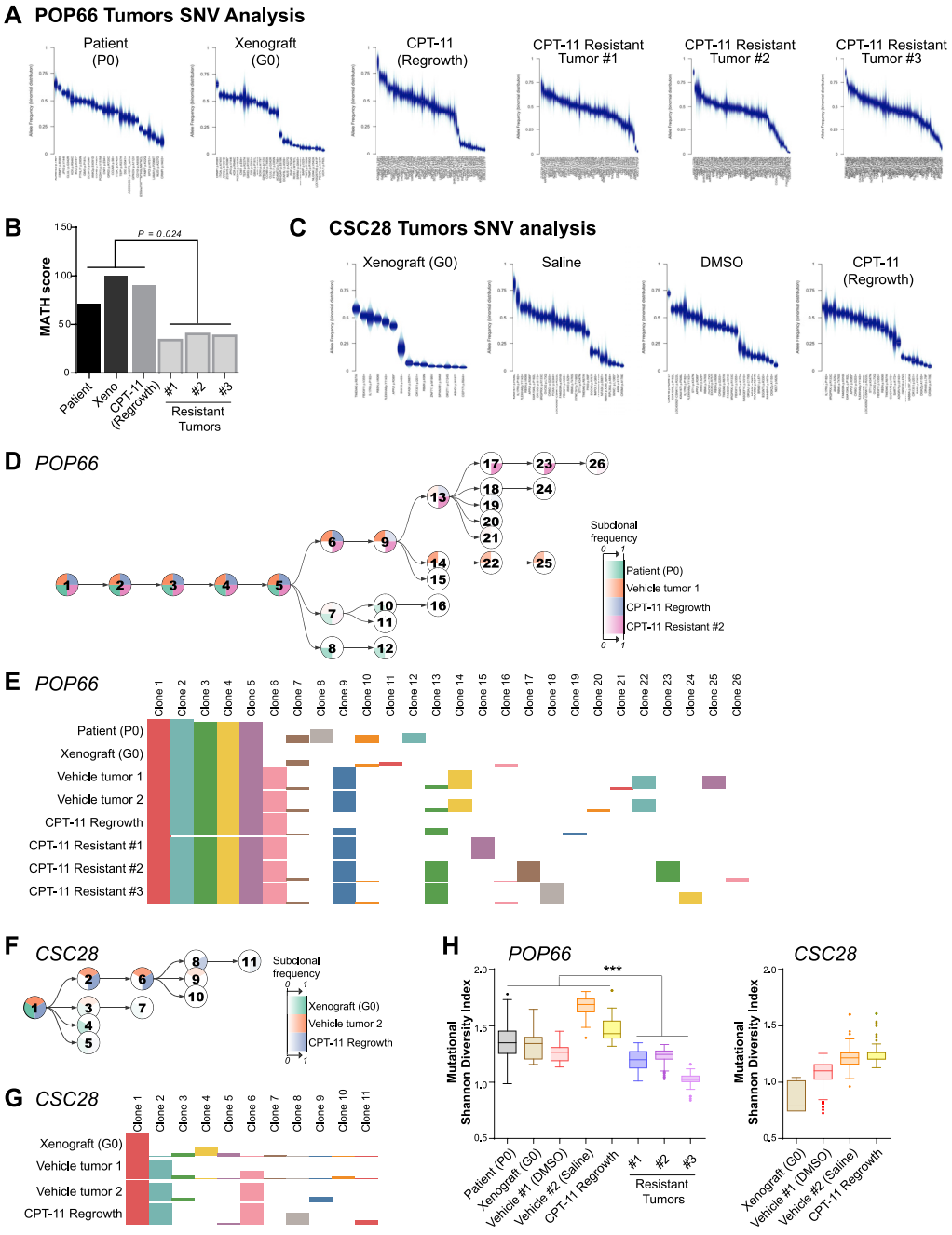


Figure S3. Genetic Intratumoral Heterogeneity Analysis, Related to Figures 1, S1, and S2 and Data S1

SNV analysis of POP66 and CSC28 tumor samples. A: Binomial distribution plots of POP66 tumor samples for each mutation in a diploid region of the genome across the full range of possible variant allele fractions (0 to 1 by 0.001 increments) as a function of sequencing depth and allele read counts. B: MATH scores plotted for the indicated POP66 samples. Statistical significance determined by unpaired t test. C: Binomial distribution plots for the indicated CSC28 tumor samples. D and E: Clone tree (D) and subclone frequencies (E) for POP66, based on SNVs in genomic regions that were diploid across all tumor samples (P0, G0, Vehicle tumors, CPT-11 Regrowth and the Resistant tumors). F and G: Clone tree (F) and subclone frequencies (G) for CSC28, based on SNVs in genomic regions that were diploid across all tumor samples (G0, Vehicle tumors and CPT-11 Regrowth). Each tree node corresponds to a genetically distinct cell subpopulation, while tree edges represent evolutionary descent. Subclone frequencies represent the proportion of cells that arose from a subpopulation and its descendants. All descendent subpopulations inherit their ancestors' mutations, such that subclone frequencies also indicate the proportion of cells possessing the mutations that first arose in the corresponding subpopulation. H: Mutational Shannon diversity indices (MSDIs) for POP66 and CSC28 samples. Statistical significance determined by Mann-Whitney rank-sum test. Box mid-lines show median; box boundaries show first and third quartiles; box whiskers show most extreme points lying within 1.5 times the interquartile range below the first quartile, and 1.5 times the interquartile range above the third quartile. ***$p < 0.001$.

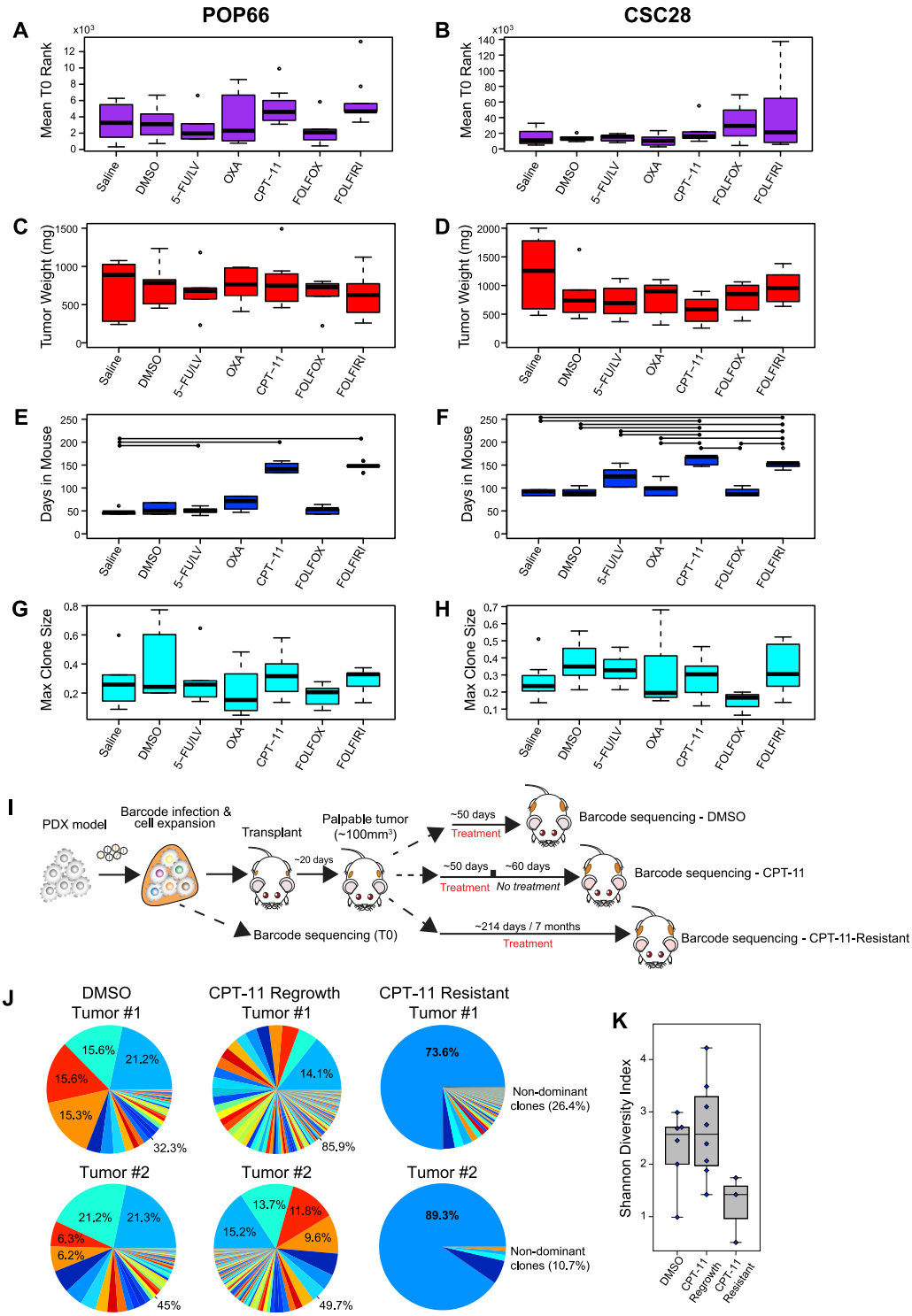


Figure S4. Characterization of CRC PDXs Post-chemotherapy, Related to Figure 2 and Table S2

A and B: Mean T0 rank of the top 10 most-enriched barcodes from each sample in each of the treatment groups. C and D: Tumor weights. E and F: Number of days from tumor implantation to sacrifice. Significant pairwise comparisons ($q < 0.05$, pairwise Wilcoxon rank-sum test with Holm multiple testing correction) indicated by horizontal lines above boxes. G and H: Distribution of the largest clones in each treatment group. A-H: Left panel plots are POP66 and right panel plots are CSC28. I: Schematic of *in vivo* barcoding and long-term treatment for the POP66 CRC PDX model from Figures 1 and S2. J: Pie charts showing barcode clone composition in POP66 CPT-11 Resistant Tumors #1 and #2. Barcode clone composition in representative (two each) DMSO and CPT-11 Regrowth tumors are also shown for comparison. K: Shannon Diversity Index plotted for barcodes from DMSO, CPT-11Regrowth and Resistant tumors for POP66. Differences in barcode composition was not statistically significant using pairwise t test.

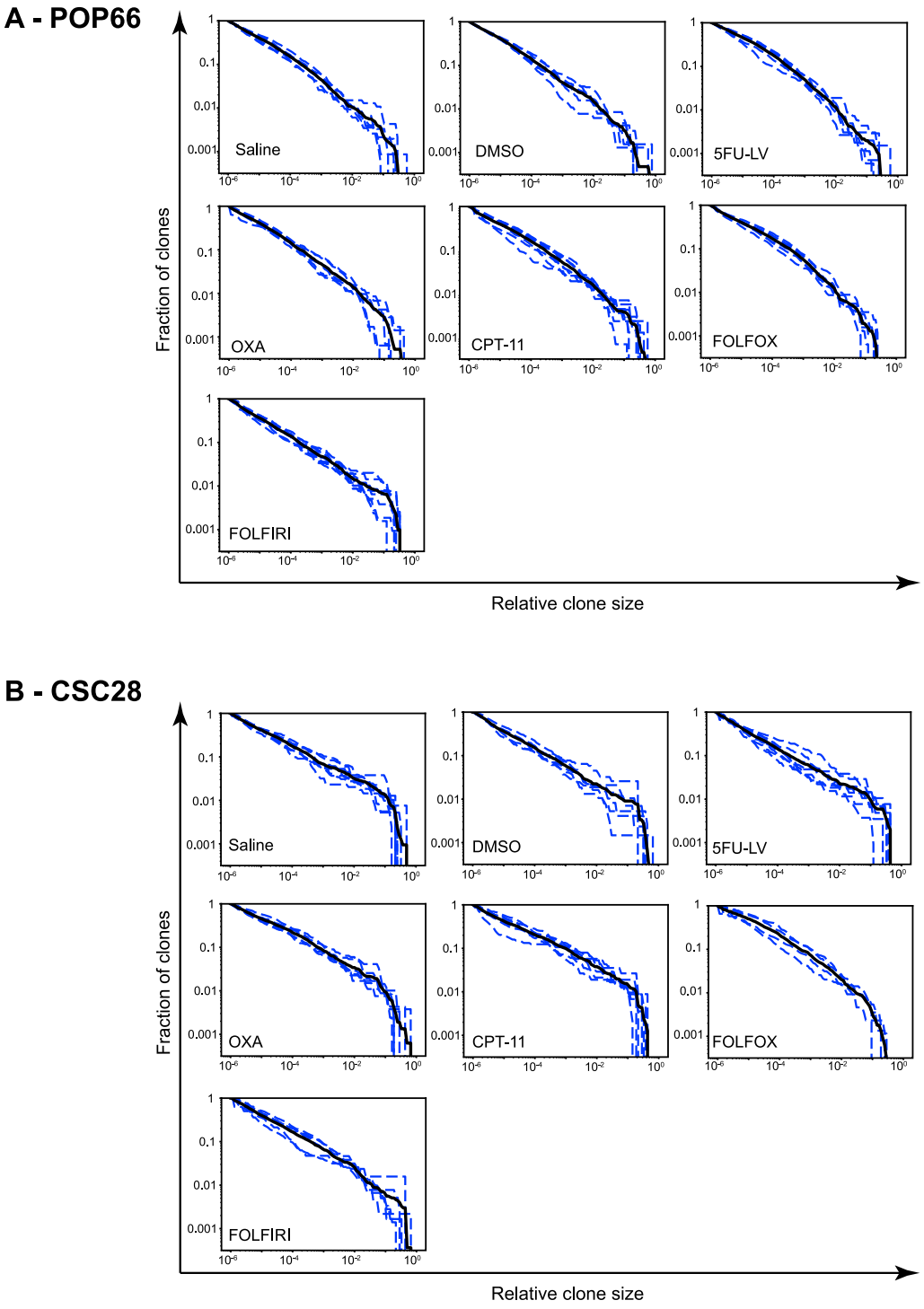


Figure S5. Clonal Composition for CRC PDXs Post-chemotherapy, Related to Figure 3

A and B: Clonal composition for POP66 (A) and CSC28 (B) xenografts under different treatments show the log-linear relationship. For each plot, dashed blue lines represent individual tumors and solid black line represents the average across all the tumors in a treatment group.

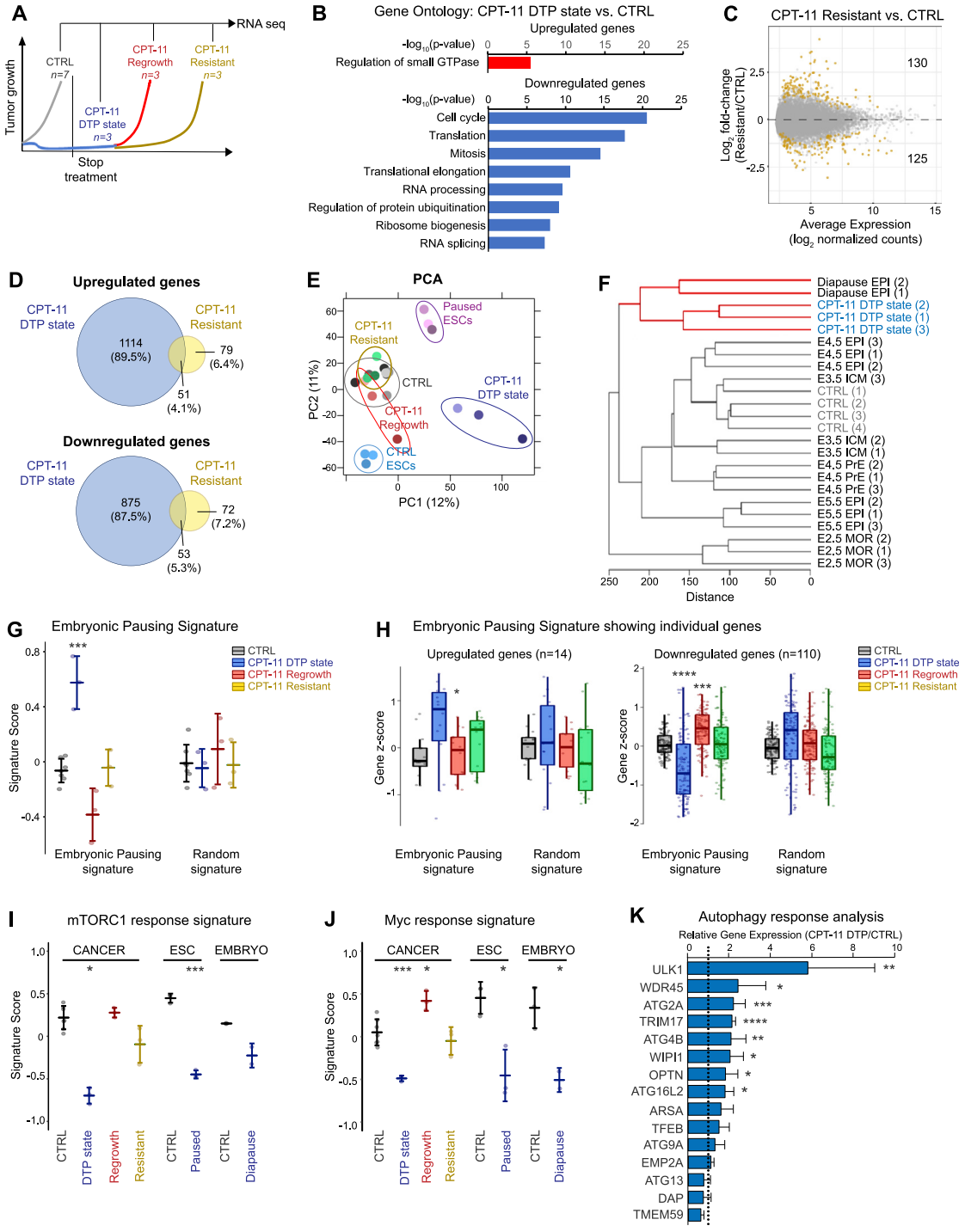


Figure S6. DTP-State Cells Express Signatures of Embryonic Pausing and Associated Signaling Pathways, Related to Figure 4 and Tables S3, S4, S5, and S6

A: Schematic diagram of the experimental protocol and sample collection for RNA-seq. B: DAVID analysis of the major gene ontology (GO) categories associated with upregulated (top panel) and downregulated (bottom panel) genes in DTP state tumors, with their respective p values ($-\log_{10}$). A curated list of GO categories (FDR < 0.05) is displayed. C: MA plot like in Figure 4B, showing $\log_2$ fold-changes in expression in CPT-11 Resistant tumor samples over CTRL. Genes differentially expressed (FDR < 0.05 and fold-change > 1.5 or < 2/3) are highlighted in green. D: Venn diagrams showing number of genes upregulated (top panel) or downregulated (bottom panel) and any overlap in genes between CPT-11 DTP-state tumors and CPT-11 Resistant tumors. E: PCA plot like in Figure 4E, including CPT-11 Resistant tumors. F: Hierarchical clustering of DMSO-treated (CTRL) tumors, CPT-11 DTP-state tumors and early embryonic developmental

(legend continued on next page)

stages, including diapaused epiblast. EPI, epiblast; MOR, morula; ICM, inner cell mass; PrE, primitive endoderm. G: Score of the embryonic diapause signature like in Figure 4G, including CPT-11 Resistant samples. Data are mean ± SD, Welch two-sample t test. H: Expression levels (z-scores) of the individual genes from the embryonic diapause signature for all samples (CTRL, CPT-11 DTP state, CPT-11 Regrowth, CPT-11 Resistant) computed from RNA-seq data. Genes up-regulated (n = 14, left panel) and downregulated (n = 110, right panel) in both paused ESCs and diapaused embryos, or an equal number of randomly selected genes, are showed separately for each sample group. Statistical significance determined by Welch two-sample t test. I: Scores of the mTORC1 signature (from the MSigDB Hallmark gene set) in all samples computed from RNA-seq data. Data are mean ± SD, Welch two-sample t test. J: Scores of the Myc signature (from the MSigDB Hallmark gene set) in all samples computed from RNA-seq data. Data are mean ± SD, Welch two-sample t test. K: Gene expression analysis of previously reported diapause-related autophagy genes in CTRL and CPT-11 DTP-state samples, taken from RNA-seq data. Data are mean ± SD, normalized to the average of the CTRL group (dashed line), Student's t test. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$, ****$p < 0.0001$.

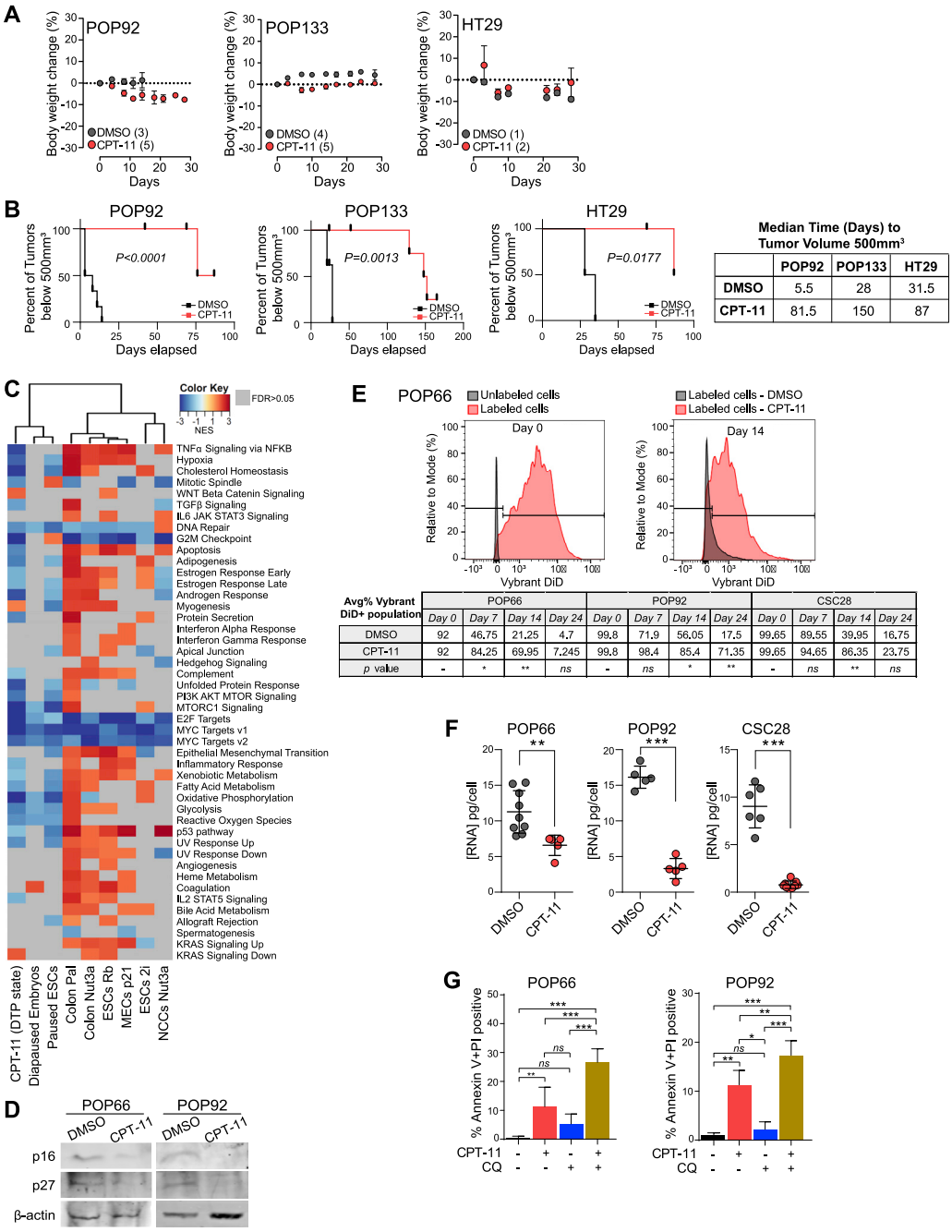


Figure S7. CPT-11-Induced Diapause-like DTP State Is Slow Cycling and Is Maintained by Upregulation of Autophagy, Related to Figures 5 and 6

A: Body weight of NOD/SCID mice bearing CRC PDXs (from Figure 5A) over the course of treatment. Body weight change was calculated as the percentage of the body weight of mice in each treatment group at day 0 of treatment. Data are shown as mean ± SEM. Numbers in parentheses denote the biological replicates in that group. B: Kaplan-Meier survival curves for CRC PDXs in Figure 5A showing tumors less than 500 mm^3 after treatment, with median time to reach 500 mm^3 for DMSO or CPT-11 treated tumors listed, log rank test. C: GSEA Hallmarks analysis of the CPT-11-induced DTP state, diapaused embryo and paused ESC models and several cell cycle arrest models as measured by the Normalized Enrichment Score (NES) for all significant pathways (FDR < 0.05). Cell cycle arrest models include: ESCs 2i: Embryonic Stem Cells cultured in MEK/GSK inhibitor containing media; Colon Nut3a:Nutlin3a treated HCT116 cells; NCC Nut 3a: Nutlin3a treated Neural Crest cells; Colon Pal: Palbociclib treated DLD1 cells; MECs p21: spontaneous p21 arrested human mammary epithelial cell line MCF10A; ESCs Rb: Rb overexpression arrested Embryonic Stem Cells. D: Western blot analysis of CDK inhibitors p16 and p27 in POP66 and POP92 cultures treated with DMSO (vehicle control) or CPT-11 (1 μM) for 14 days *in vitro*. Data shown are representative of at least n = 3 independent experiments. E: Representative cytofluorimetric histograms of Vybrant DiD labeling for POP66 cultures at day 0 (left plot) or on day 14 of DMSO and CPT-11 (1 μM; right plot) treatment (for data in Figure 5G).

(legend continued on next page)

Table indicates cytofluorometrically determined percent of Vybrant DiD label positive cultures treated with DMSO or CPT-11 (1 μM) at 0, 7, 14 and 24 days after the start of treatment. Treatment was stopped on day 14. F: Total RNA amount per cell in POP66, POP92 and CSC28 cultures treated with DMSO or CPT-11 (1 μM) for 14 days. Data shown as mean ± SD of at least n = 3 independent experiments, t test. G: Apoptosis/Necrosis analysis (Annexin V/PI label) performed on POP66 and POP92 cultures treated with DMSO, CPT-11 (1 μM), Chloroquine (CQ, 20 μM) or combination for 5 days. Data shown as mean ± SD, n = 3 independent experiments, one-way ANOVA. *p < 0.05, **p < 0.01, ***p < 0.001, *ns* = not significant.