Title
Non-genetic factors influencing drug persistence and survival in cancer cells

Author
Morinishi, Leanna

Publication Date
2020

Peer reviewed|Thesis/dissertation

Non-genetic factors influencing drug persistence and survival in cancer cells

by
Leanna Morinishi

DISSERTATION
Submitted in partial satisfaction of the requirements for degree of
DOCTOR OF PHILOSOPHY

in

Biological and Medical Informatics

in the

GRADUATE DIVISION
of the
UNIVERSITY OF CALIFORNIA, SAN FRANCISCO

Approved:

DocuSigned by:
Steven J. Altschuler
3BE37B146DB5437...

Steven J. Altschuler
Chair

DocuSigned by:
Lani F. Wu
[illegible]4BF...

Lani F. Wu

DocuSigned by:
Michael McManus
[illegible]455...

Michael McManus

DocuSigned by:
Alexander Marson
7F25CDCE383C4A8...

Alexander Marson

Committee Members

Acknowledgements

First and foremost, I'd like to thank my family for their unconditional support and enthusiasm for my goals and efforts. To my parents, Ron & Carolyn, you have been my main source of inspiration throughout my life. Thank you for sharing your love of science with me and encouraging my own. I'll never forget looking at planets with the telescope in the park or feeling the solid weight of my very own microscope. Thank you for working so hard to make sure that the world was open to me, that I felt safe and brave enough to pursue my every interest. I love you, and I truly appreciate all that you've done. To my siblings Justin and Melissa, you two are the kindest, funniest, most talented humans I've ever known. You have grown up to be truly amazing people, and I'm fortunate to have you as role models inspiring me to be the best person I can be. To my cousins, aunts and uncles, and especially to my Grammy Grace and my 'Baachan Marian, thank you for your patient love and support, thank you for always believing in me. To my grandfathers Bob Yoshihiro and Mike Yoshio, both scientists in their own right, I only hope that I've made you proud.

Over the past six years my life has transformed in the best possible ways, including newly acquired family. To the Hodges, the Goodhand clan, and the Fraimows, thank you so much for making me feel loved and welcomed. Lynne & Josh, I truly appreciate your patience, kindness, and support throughout the years. To my new sister Jocelyn, you are a joy and a superstar, and I'm lucky and grateful to call you family.

I've also been blessed with friends in and out of the lab that have made these past years great. Thanks to Lia, Roxana, and Jillia of Troop 452 for being incredible sisters to me for the past 20+ years! It's been truly special to blindly fumble my way through adulthood with you, and I'm excited to see what our future holds. A special and heartfelt thank you to my recruitment roomie and biffle Elena Cáceres who was always down for coffee or a quarantine phone call, and who married me… to my husband. My grad school experience would have been so much sadder and less caffeinated without you. You're an incredible, generous, genuine person and you constantly inspire me to be a better human and a better scientist. <3 Special thanks to Anime Knights for a crash course in amazing shows and movies, for memorable sluricane karaoke, and for giving me the space to scream into the void. Thank you to all my former roommates (and

their pets) for making our apartments feel like home, and a big shout-out to Brunch Crew for giving me an excuse to get outside Saturday mornings.

I started this journey to learn scientific rigor and become an independent thinker. This would not have been accomplished without my mentors Steven Altschuler and Lani Wu. Over the years, they have given me both the intellectual freedom to pursue my interests and the guidance necessary to develop raw data into a scientific story. They were always generous with their time and open with their feedback, and their mentorship has certainly shaped the scientist I am today. The Altschuler and Wu Labs has been a great place to call home these past years, and I have learned so from each and every person. In particular, I want to thank my coauthor Karl Kochanowski for his scientific guidance, thoughtful advice, and good humor. To Wetlab Crew past and present, especially Louise Heinrich, Laura Sanman, Maike Roth, and Xiaoxiao Sun (and sometimes Jeremy Chang) thank you for making my days in the wetlab fun and full of avocados. Sometimes I even forgot about the roaring air filter and the constantly beeping liquid nitrogen tank. To my fellow AWlab iPQBers Ina Chen and Weiyue Ji, I know grad school was never easy, but I'm so glad we were able to undertake these proverbial growing pains together.

I also want to thank former mentors who have given me the opportunity to do research, and people who have helped me learn and grow. Thank you to my committee members Michael McManus and Alex Marson whose scientific and professional guidance has been essential in my training. Thank you as well to my former mentors Paul Blainey, Ron Weiss, James Heath, Al Grodzinsky, Deepak Mishra, Gwen Owens, Chao Ma, and Emily Florine. I truly appreciate the time you took to train me and the encouragement you gave me to continue with science. Thanks as well to my mentees Janice Goh, Eva Hansen, Xiaofei, Christine, Dana, Arohee, and Erica. It was a true privilege (and a learning experience) watching you grow as scientists.

And finally, none of this would have been possible without the patience, support, and understanding of my partner Izzy. I cannot put into words how much you mean to me, and I cannot thank you enough for all you have done for me. Your principles, humor, honesty, and resolve have changed me forever for the better, just one of the countless reasons why you're the best. I'm so glad to have you in my life and I am beyond excited to begin the next stage of our lives together.

Acknowledgements of previously published materials and research contributions

Chapter 1 of this dissertation was previously published. The text in Chapter 1 is a reprint and expansion of the materials as it appears in Current Opinion in Systems Biology (Kochanowski & Morinishi *et al*. 2018). Lani F. Wu and Steven J. Altschuler supervised the research that forms the basis of this chapter. Karl Kochanowski and Leanna Morinishi wrote the manuscript and created all tables and figures.

Chapter 2 of this dissertation was previously published. The text in Chapter 2 is a reprint and expansion of the materials as it appears on bioRxiv and in Cell Systems (Morinishi *et al*. 2020). Karl Kochanowski was involved with data generation. Ross L. Levine contributed key resources. Lani F. Wu, Steven J. Altschuler, Karl Kochanowski and Ross L. Levine supervised the research that forms the basis of this chapter and were involved in the writing process. The primary data generation and analysis was performed by Leanna Morinishi who also wrote the manuscript and created all tables and figures.

The contents of Chapter 3 are unpublished. Karl Kochanowski, Maike Roth, and Maren Diether contributed data and analysis to this work. Janice Goh and Mollie Hansen were involved with data generation. Figures 3.2-3.4, and Tables 3.1-3.2 were created by and used with permission of Karl Kochanowski. Lani F. Wu, Steven J. Altschuler, and Karl Kochanowski supervised the research that forms the basis of this chapter. Leanna Morinishi wrote the chapter and created all other tables and figures.

Contributions

Abstract:

Ramirez, M; Rajaram, S; Steininger, FJ; Osipchuk, D; Roth, MA; Morinishi, LS; Evans, L; Ji, W; Hsu, CH; Thurley, K; Wei, S; Zhou, A; Koduru, PR; Posner, BA; Wu, LF & Altschuler, SJ. Diverse drug-resistance mechanisms can emerge from drug-tolerant cancer persister cells. Nature Communications, 2016.

Chapter 1:

Kochanowski, K & Morinishi, L; Altschuler, SJ & Wu, LF. Drug persistence – From antibiotics to cancer therapies. Current Opinion in Systems Biology, 2018.

Chapter 2:

Morinishi, L; Kochanowski, K; Levine, RL; Wu, LF & Altschuler SJ. Loss of TET2 affects proliferation and drug sensitivity through altered dynamics of cell-state transitions. Cell Systems, 2020.

Non-genetic factors influencing drug persistence and survival in cancer cells

Leanna Morinishi

Abstract

Despite the development of new, promising therapies for treating cancer, the emergence of cancer cells that do not respond to treatment remains a major barrier to cure. These emerging populations of cancer cells can result from secondary mutations that give rise to inherited drug tolerance or resistance, but in many cases there is no clear genetic basis for improved drug survival. In fact, prior work has shown that non-genetic mechanisms (e.g. chromatin remodeling) can enable longer-term persistence of cells in drug (Sharma *et al*., 2010, Hinohara *et al*., 2018), and we and others have demonstrated that these "persister" cells are an important reservoir for the emergence of drug-resistant mutants (Ramirez *et al*., 2016, Hata *et al*., 2016). Here through a combination of data analysis, mathematical modeling, and borrowed insight from microbiology, we further our understanding of non-genetic factors that alter cancer cell response to drug treatment.

In the first chapter, I summarize literature describing the persister state in both antibiotic-treated bacteria and drug-treated cancer cells, and I highlight shared features of these persisters (such as slow growth and distinct metabolic activity). Of particular relevance to this dissertation, I discuss epigenetic and metabolic changes that influence cancer progression and drug survival. In the second chapter, I pair a close study of cancer cells differing only by expression of the epigenetic modifier TET2 with mathematical modeling to demonstrate how TET2 loss can confer a fitness advantage in drug by altering the dynamics of cell-state switching. Finally, in the third chapter, I present unpublished work focused on metabolic changes and vulnerabilities associated with TET2 loss. Through these studies, I describe several non-genetic mechanisms that affect persistence and tolerance of cancer cells in drug and highlight the challenges in pharmacologically targeting these complex and ill-defined modes of drug survival.

Table of contents

List of Figures

List of Tables

List of Abbreviations

AML: acute myeloid leukemia

CD: cluster of differentiation

CpG: shorthand for single-stranded linear sequence 5’—C—phosphate—G—3’

CSC: cancer stem cell

DNA: deoxyribonucleic acid

EGFR: epidermal growth factor receptor

FACS: fluorescence-activated cell sorting

MRD: minimal residual disease

mRNA: messenger RNA

PCR: polymerase chain reaction

(p)ppGpp: guanosine pentaphosphate or tetraphosphate

RNA: ribonucleic acid

RT-qPCR: reverse transcription followed by quantitative polymerase chain reaction

scRNAseq: single-cell RNA sequencing

TET2: tet methylcytosine dioxygenase 2

Chapter 1

Drug persistence - from antibiotics to cancer therapeutics

Main contributing authors

Leanna Morinishi[1,*], Karl Kochanowski[2,*], Lani F. Wu[2] & Steven J. Altschuler[2]

Affiliations

[1]Bioinformatics Graduate Group, University of California, San Francisco

[2]Department of Pharmaceutical Chemistry, University of California, San Francisco

[*]Equal contribution

Abstract

Drug-insensitive tumor subpopulations remain a significant barrier to effective cancer treatment. Recent works suggest that within isogenic drug-sensitive cancer populations, subsets of cells can enter a 'persister' state allowing them to survive prolonged drug treatment. Such persisters are well-described in antibiotic-treated bacterial populations. In this review, we compare mechanisms of drug persistence in bacteria and cancer. Both bacterial and cancer persisters are associated with slow-growing phenotypes, are metabolically distinct from non-persisters, and depend on the activation of specific regulatory programs. Moreover, evidence suggests that bacterial and cancer persisters are an important reservoir for the emergence of drug-resistant mutants. The emerging parallels between persistence in bacteria and cancer can guide efforts to untangle mechanistic links between growth, metabolism, and cellular regulation, and reveal exploitable therapeutic vulnerabilities.

Introduction

The last decades have brought the arrival of an impressive arsenal of therapies for treating cancer. At the same time, countless drug resistance mechanisms have been discovered by which cancer cells avoid and subvert drug treatment. Tumor subpopulations that do not respond to therapeutics are a significant barrier in the treatment of cancer, and cancer remains a major global killer (Torre *et al*., 2012).

Recently, it has become clear that even within otherwise drug-susceptible isogenic cancer populations, a subset of cells can enter a persister state, in which they survive prolonged drug exposure (Sharma *et al*., 2010; see Table 1.1 for a list of cancer persister models). While this persister state has only recently started to draw attention in mammalian cells, bacterial persisters were described in literature as early as 70 years ago (Bigger, 1944). The past decade has seen a surge in studies elucidating the mechanisms underlying bacterial antibiotic persistence – as recently summarized in a string of excellent reviews (Fisher *et al*., 2017; Brauner *et al*., 2016; Radzikowski *et al*., 2017; van den Bergh *et al*., 2017). In this review, we compare and contrast persistence in bacteria and cancer cells and highlight surprising parallels in the underlying persistence mechanisms.

Discussion

Defining persistence – a persistent challenge

Before delving into persistence mechanisms, we must first define what drug persistence is, and how it differs from other mechanisms of drug insensitivity (Figure 1.1). Bacterial insensitivity to antibiotics is classified phenomenologically into three broad categories that can be distinguished experimentally (compared to a reference sensitive population; Figure 1.1), as summarized in Fisher *et al*., 2017 and Brauner *et al*., 2016. The first category, drug tolerance, is the ability of cell populations to withstand transient lethal antibiotic concentrations, while remaining genetically susceptible. Experimentally, tolerance manifests as a decreased rate of killing during drug exposure compared to a sensitive reference population (Figure 1.1). The second category, drug resistance, is the genetically inherited ability of cells to grow at normally lethal

antibiotic concentrations (Fisher *et al*., 2017 and Brauner *et al*., 2016). Drug-resistant populations show a characteristic increase in minimal inhibitory concentration (the lowest drug concentration needed to prevent bacterial growth); this increase is absent in drug-tolerant populations. In contrast to these two categories, which are defined at the population level, drug persistence describes scenarios in which only a subpopulation of cells within a clonal cell population survives prolonged antibiotic treatment, while remaining genetically susceptible to reapplication of the drug (Balaban *et al*., 2013). An important feature of bacterial drug persistence is its phenotypic reversibility. After drug treatment is stopped, the remaining persister cells will eventually re-establish a population showing the same heterogeneous response when re-treated with the same drug (Figure 1.1). Experimentally, drug persistence is characterized by a survival curve with two phases - an initial steep decline in cell number followed by a cell number plateau -, which is absent in drug-tolerant populations (Brauner *et al*., 2016; van den Bergh *et al*., 2017; Figure 1.1).

Compared with the converging literature view of how to define and distinguish bacterial persisters, terminology is somewhat more diverse in cancer literature. Persistence is sometimes used interchangeably with drug tolerance to describe subpopulations that have an enhanced (and non-genetic) ability to survive drug treatment (Sharma *et al*., 2010; Hata *et al*., 2016; Hangauer *et al*. 2017). Various other terms have also been used to describe scenarios in which a phenotypically distinct subpopulation survives prolonged drug treatment, including quiescence (Chen *et al*., 2012), dormancy (Viale *et al*., 2014) or cancer stem cells (Gupta *et al*., 2011; Figure 1.2). Throughout this review, we will use the term ‘persistence’ for cases in which a subpopulation survives drug treatment but regains sensitivity after drug removal, and we reserve the term ‘tolerance’ for cases in which the whole population is more resilient to drug exposure.

Paths to persistence

How do cells become persisters? We will first briefly discuss mechanisms of bacterial antibiotic persister formation, and then relate these to our current understanding of how cancer drug persisters emerge. In particular, we will focus on the impact of three factors on persistence: cell growth, metabolic activity, and regulatory program.

Arguably the best studied bacterial persistence mechanism are Toxin-Antitoxin (TA) systems (Page *et al*., 2016). These consist of a stable toxin, which arrests growth by inhibiting vital cellular processes such as transcription or translation thereby inducing the persister state, and a labile antitoxin acting as the antidote (van den Bergh *et al*., 2017). An example is the HipBA module in E. coli, which inhibits the glutamyl-tRNA synthetase GltX and thus halts translation (Germain *et al*., 2013; Kaspy *et al*., 2013). Originally identified as a mechanism to prevent plasmid loss, TA systems were shown to induce the stochastic formation of non-growing persisters in exponentially growing cultures (Balaban *et al*., 2004; Maisonneuve *et al*., 2013).

Recent works have identified additional factors that modulate antibiotic persistence. For example, various studies found that the fraction of persisters in different environmental conditions is inversely correlated with the population growth rate, as shown e.g. in (Maisonneuve *et al*., 2013; Fung *et al*., 2010) and summarized in Brauner et al, 2016. Additionally, stresses, such as salt-stress, can increase the rate of persister formation (Shan *et al*., 2017). Particularly interesting types of environmental stress are shifts in nutrient availability: bacteria undergoing nutrient shifts, which are typically accompanied by a transient reduction in growth rate, show dramatically elevated persister fractions (Amato *et al*., 2013; Kotte *et al*., 2010; Radzikowski *et al*., 2016). The examples above evoke a 'tolerance by slow growth' (Brauner *et al*., 2016) scheme, in which slow-growing bacteria tend to become more resilient against antibiotic treatment, regardless of how exactly the reduction in growth rate came about.

This increase in antibiotic persistence at slow growth could of course simply be the consequence of a reduction in the activity of the antibiotic targets, i.e. the cellular transcription/translation machinery. However, mounting evidence suggests that antibiotic persistence in fact relies on an active cellular program. Various studies have demonstrated that the (p)ppGpp-mediated bacterial starvation program (also termed "stringent response") modulates the rate of persister formation (Maisonneuve *et al*., 2013; Amato *et al*., 2015; Nguyen *et al*., 2011; Verstraeten *et al*., 2015), for example by directly inducing the expression of TA systems (Maisonneuve *et al*., 2013). Importantly, mutant strains lacking the stringent response program are

readily killed by antibiotic treatment even in starvation conditions (Nguyen *et al*., 2011), suggesting that absence of growth alone is not sufficient to induce persistence.

Finally, recent studies have shown that persisters can be selectively killed off by modulating their metabolic activity (Shan *et al*., 2017; Nguyen *et al*., 2011; Allison *et al*., 2011; Meylan *et al*., 2017). For example, addition of metabolic stimuli that promote an increase in proton-motive force by the oxidative electron transport chain triggers the uptake of aminoglycoside antibiotics in persister cells, thereby enhancing persister killing (Allison *et al*., 2011). These observations indicate that persisters retain a distinct - and active - metabolic state, and highlight the importance of metabolism for persistence (Radzikowski *et al*., 2017; Amato *et al*., 2014).

Collectively, these studies paint a picture of bacterial antibiotic persistence with three main themes. The first theme is the recurrent observation that slow bacterial growth tends to favor persister formation. The second theme is a distinct metabolic state in persisters, which leaves them vulnerable to metabolic perturbations. The third theme is the reliance on a specific regulatory program involving the stringent response, which can be directly targeted to reduce persister formation.

Do these themes have parallels in cancer drug persisters? Recent literature suggests that this is indeed the case. The first hint stems from the observation that slow-growing cancer cells also tend to be more drug-tolerant in a wide range of cell types and model systems (Chen *et al*., 2012; Fallahi-Sichani *et al*., 2017; Su *et al*., 2017; Liau *et al*., 2017; Roesch *et al*., 2010; Jordan *et al*., 2016; Roesch *et al*., 2013). Such drug-tolerance has been reported during exposure to chemo- (Roesch *et al*., 2013) and targeted (Fallahi-Sichani *et al*., 2017) therapy, for adherent (Liau *et al*., 2017) and suspension (Su *et al*., 2017) cells *in vitro*, as well as *in vivo* in mouse models (Chen *et al*., 2012), suggesting a general phenomenon.

Surprisingly, there is also evidence that metabolism might be an important determinant of drug persistence in cancer cells (Hangauer *et al*., 2017; Viale *et al*., 2014; Roesch *et al*., 2013; Raha *et al*., 2014; Viswanathan *et al*., 2017; Viale *et al*., 2016). Recent works by Hangauer *et al*. and Viswanathan *et al*. demonstrated that various persister models are vulnerable to inhibition of the lipid hydroperoxidase GPX4 (Hangauer *et al*., 2017; Viswanathan *et al*., 2017). GPX4 catalyzes the glutathione-dependent reduction of

lipid peroxides, which cause oxidative stress, and thereby prevents the induction of ferroptosis, a non-apoptotic form of cell death (Dixon *et al*., 2012). This result supports earlier work showing that the enzyme aldehyde dehydrogenase 1 A1 (ALDH1A1)—which is also involved in lipid peroxidation and expressed in many cancer stem cells—is required to maintain drug persistence (Raha *et al*., 2014). Moreover, recent reports indicate that persisters rely more heavily on oxidative phosphorylation (OXPHOS), and are more sensitive to OXPHOS inhibitors (Viale *et al*., 2014; Roesch *et al*., 2013; Wolf, 2014). In particular, these subpopulations were reported to have a diminished 'glycolytic reserve', which is the ability to increase glucose uptake for ATP generation if OXPHOS is inhibited, suggesting impaired metabolic plasticity (Viale *et al*., 2014). Whether these changes in metabolic activity are an adaptation to redox stress (Sabharwal *et al*., 2014; Gorrini *et al*., 2013), or rather reflect an increased demand for ATP (Viale *et al*., 2016), is currently unclear.

Finally, several lines of evidence suggest that drug persistence in cancer cells also relies on a distinct regulatory program (Inde *et al*., 2017), and particularly pinpoint two regulatory processes. The first process is epithelial-to-mesenchymal transition (EMT), which causes cells to gradually lose their differentiation status and become more stem-cell like (Kreso *et al*., 2014). Expression of EMT/stem-cell markers is a frequent hallmark of persister subpopulations, which can be exploited to isolate persister cells within isogenic populations (Fallahi-Sichani *et al*., 2017; Su *et al*., 2017; Roesch *et al*., 2013; Raha *et al*., 2014; Pisco *et al*., 2013). Second, several lines of evidence point towards chromatin remodeling as a key step in persister formation (Sharma *et al*., 2010; Liau *et al*., 2017; Vinogradova *et al*., 2016; Guler *et al*., 2017). For example, inhibition of the histone demethylases KDM5 and KDM6 were found to suppress the emergence of persisters (Roesch *et al*., 2013; Vinogradova *et al*., 2016). These findings are particularly intriguing given that epigenetic and metabolic changes seem to be closely linked in many cancers. A prime example are mutations in metabolic enzymes, such as isocitrate dehydrogenase 1 and 2, which can modulate the epigenetic state of cells through the accumulation of 'oncometabolites' such as 2-hydroxyglutarate, and thereby influence cancer progression and drug survival (Mingay *et al*., 2017; Flavahan *et al*., 2016). The involvement of EMT and chromatin remodeling suggests the requirement for a distinct regulatory program

to ensure the formation and/or maintenance of a persister state, analogous to the aforementioned stringent response dependent regulation in bacterial persisters. As we will discuss in more detail in the final section of the review, a major open question is how these processes are linked mechanistically in cancer persisters.

Despite these striking parallels, there are also bacterial persistence mechanisms that do not have a direct analog in cancer. For example, a study found that E. coli treated intermittently with Ampicillin for different durations quickly evolved population lag times (the time it takes to transition from stationary to exponential growth phase) to match the duration of the antibiotic exposure, while the maximal growth rate did not change (Fridman *et al*., 2014). To our knowledge, there is only one example of such 'tolerance by lag' (Brauner *et al*., 2016) in the cancer literature (Pearl *et al*., 2016). Nevertheless, since cancer therapies often involve drug administration at regular time intervals, it is at least conceivable that similar selective pressures might also affect cancer cell populations.

Conversely, there are also cancer persistence mechanisms with no direct bacterial equivalent. A compelling example was recently presented by Shaffer *et al*., 2017. In an elegant set of experiments in patient-derived melanoma cell lines, the authors demonstrated the existence of transiently pre-resistant subpopulations characterized by sporadic expression of resistance markers, for example alternative oncogenes such as EGFR, which can become more tolerant of a given targeted therapy. Conceptually, such 'tolerance by sporadic bypassing' is similar to the sporadic high expression of multidrug efflux pumps in bacterial populations (Pu *et al*., 2016). Another similar phenomenon is the aforementioned stochastic activation of TA systems – and subsequent switch to a slow/non-growing cell state – in exponentially growing bacterial cultures. However, it is currently not clear whether these pre-resistant cancer subpopulations also adopt a non-growing state. Another open question is whether this mechanism constitutes a 'bug or feature' of mammalian regulatory networks: is the sporadic activation of signaling kinases merely the inevitable consequence of stochastic fluctuations within highly nonlinear signaling networks, or an evolved bet-hedging strategy? As we will discuss in the next section, recent evidence suggests that at least bacterial persistence may indeed be an evolvable trait.

Persistence and evolution

Given the ubiquitous cell-to-cell variability in gene expression, it is tempting to assume that persistence is an inevitable byproduct of life ('persistence as stuff happens') (Levin *et al*., 2014). However, there is some evidence that bacterial persistence is actually an evolvable trait (van den Bergh *et al*., 2016; Mechler *et al*., 2015). E. coli cultures exposed to different frequencies of antibiotic treatment quickly evolve an inverse relationship between persister fractions and treatment interval, without altering their antibiotic resistance (van den Bergh *et al*., 2016). This observation is of particular importance given evidence that antibiotic tolerance acts as a stepping-stone on the path to resistance. A study by Levin-Reisman *et al*. demonstrated in a series of *in vitro* evolution experiments that Ampicillin-resistant E. coli mutants emerge from the pool of initially only Amp-tolerant mutants (Levin-Reisman *et al*., 2017). A potential explanation for this result is that the space of mutations conferring tolerance – and therefore the probability to establish a tolerance-inducing mutation – is substantially larger than the space of mutations conferring resistance (Brauner *et al*., 2016).

Interestingly, recent work suggests that cancer persister cells are also an important reservoir for the emergence of resistant cell lines *in vitro* (Hata *et al*., 2016; Ramirez *et al*., 2016). Hata *et al*. focused on the clinically relevant EGFR-T790M gatekeeper mutation, which makes EGFR-mutant non-small-cell lung cancer (NSCLC) cells resistant to EGFR inhibitors. The authors found that EGFR-T790M positive populations not only originate from the selection of pre-existing mutants, but they can also emerge from the pool of persister cells. Moreover, work by Ramirez *et al*. suggests that even within an initially isogenic EGFR-mutant NSCLC cell population exposed to an EGFR-inhibitor, the persister subpopulation ultimately gives rise to different mutant populations with diverse resistance mechanisms (Ramirez *et al*., 2016). These results indicate that the evolution of drug resistance is not necessarily restricted towards few attainable bypass mechanisms when cell populations first pass through a persister state.

If resistant mutants indeed evolve from the pool of persister cells in a population, a rational strategy to minimize their emergence is the elimination of persisters before drug exposure. However, as we discussed above, persisters not only emerge spontaneously in untreated populations, but can also be induced

by various environmental stresses. Recent *in vitro* studies have begun to elucidate the importance of induced persistence, also termed 'type I persistence' (Balaban *et al*., 2013), in cancer cell populations (Su *et al*., 2017; Pisco *et al*., 2013). Pisco *et al*. showed in clonally derived leukemia cells that the rapid emergence of multidrug resistance 1 (MDR1) mediated persisters upon chemotherapeutic treatment is largely driven by induced persistence (Pisco *et al*., 2013). Whether such 'Larmarckian induction' is the exception or the norm in the emergence of cancer drug persisters remains an open question. Nevertheless, these works suggest that inhibition of the mechanisms that mediate the transition to a persister state during drug treatment may help to prevent the emergence of drug resistance in cancer (Su *et al*., 2017; Pisco *et al*., 2013; Ramirez *et al*., 2016; Pisco *et al*., 2015; Goldman *et al*., 2015).

Conclusions and open questions

In this review, we explored the surprising parallels between bacterial and cancer persisters that are emerging in recent literature. In particular, both persister types are frequently associated with a slow-growing phenotype, show metabolic alterations that leave them vulnerable to metabolic perturbations, and rely on a distinct regulatory program that can be targeted to prevent persister formation.

Currently, the mechanistic links between the emerging regulatory processes in cancer persisters, namely EMT and chromatin remodeling, their slow-growth phenotype, and their vulnerability to inhibitors of lipid peroxidation and oxidative phosphorylation (Hangauer *et al*., 2017; Viale *et al*., 2014; Roesch *et al*., 2013; Raha *et al*., 2014; Viswanathan *et al*., 2017), are unclear. Lessons from bacterial research may provide some inspiration. For example, recent works have shown that the global coordination of protein expression in bacteria heavily depends on the growth rate (Scott *et al*., 2010; Hui *et al*., 2015; Schmidt *et al*., 2016; Borkowski *et al*., 2016; Kochanowski *et al*., 2017) and can be described by few so-called 'growth laws' (Scott *et al*., 2011). It is tempting to speculate that in cancer cells similar mechanisms could potentially induce an EMT-type transcriptional program if growth is impaired. Such a mechanism might explain the rapid increase in persister fraction that has been observed in drug exposed populations (Su *et al*., 2017; Pisco *et al*., 2013). Another intriguing question is whether metabolism can directly induce

persistence in cancer cells, similar to the nutrient-shift induced persisters in bacteria (Amato *et al*., 2013; Kotte *et al*., 2010; Radzikowski *et al*., 2016; Amato *et al*., 2015). Interestingly, recent reports suggest that loss of fumarate hydratase may induce EMT through fumarate-mediated changes in epigenetic state (Sciacovelli *et al*., 2016), thus providing a potential link from metabolism to EMT-induced drug tolerance. Future efforts might identify scenarios in which metabolism drives cancer persister formation.

A major aspect not discussed in this review is the influence of the microenvironment on persister formation. Bacteria living in communities—termed biofilms—tend to have higher persister fractions than planktonic cultures (Harms *et al*., 2016). Do nearby cells also affect the formation of cancer persisters, for example through direct cell contact or indirectly via signaling molecules? *in vitro* observations have shown that growth factors can attenuate the efficacy of oncogene-targeting drug therapy by activating alternative signaling pathways (Wilson *et al*., 2012). Future studies might focus on identifying additional tumor microenvironment signals that play a role in the formation and maintenance of cancer persisters.

Finally, the relevance of persisters in clinical settings remains an open question. There is indeed evidence that antibiotic persisters play a role in bacterial infections (Fisher *et al*., 2017). For example, Pseudomonas aeruginosa strains infecting patients with cystic fibrosis show dramatically increased persister levels over time, which seems to be the main mechanism to cope with antibiotic treatment (Mulcahy *et al*., 2010). There is also evidence in murine Salmonella typhimurium infections that slow-growing persisters survive antibiotic treatment and drive disease progression (Claudi *et al*., 2014). In contrast, the role of persisters in tumor progression is more enigmatic. So far, evidence is mostly restricted to mouse models (Chen *et al*., 2012; Kreso *et al*., 2013). For example, slow-growing glioblastoma subpopulations were reported to survive initial drug exposure and repopulate the tumor after cessation of drug treatment (Chen *et al*., 2012). It is currently unclear which of the aforementioned persister-targeting strategies – exploiting the distinct metabolic vulnerabilities of cancer persisters or preventing the transition into a persister state in the first place – will be successful in clinics. An important step in the development of such therapies will be the identification of molecular signatures that are unique for persisters, enabling the detection of cancer persisters in clinical samples throughout treatment.

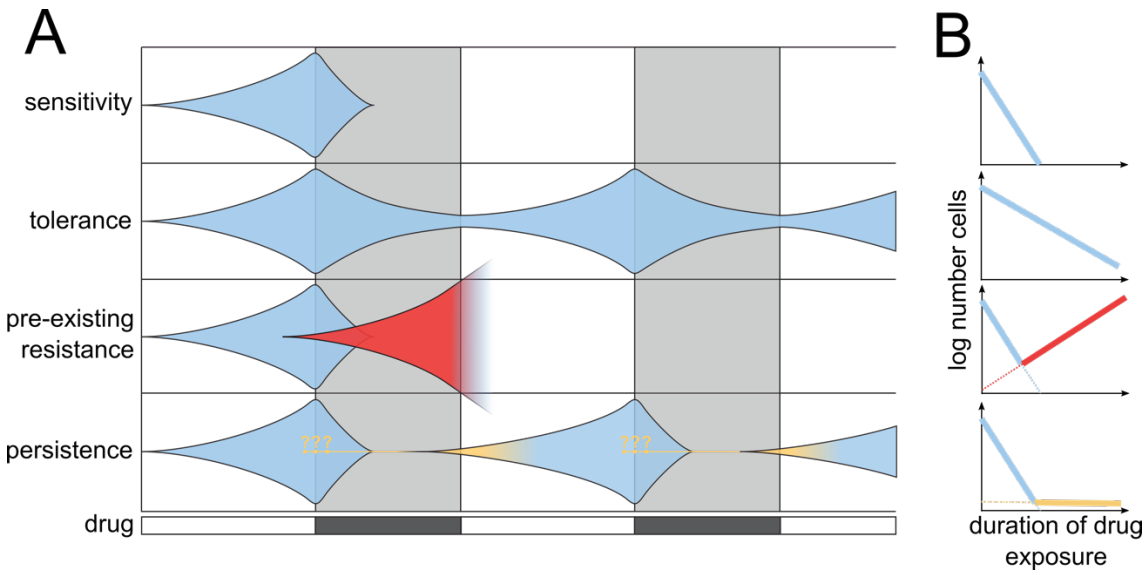

Figure 1.1: Schematic of different forms of drug sensitivity
(A) Blue area: Size of cell population. Grey shaded area: duration of drug treatment. Sensitivity: cells in a drug-sensitive population are readily killed by the drug. Tolerance: a drug-tolerant population is killed at a slower rate than a sensitive population. Pre-existing resistance: drug-treatment selects pre-existing resistant mutants (red), which continue growing while sensitive cells are being killed. Persistence: persister subpopulations (yellow) form either before drug treatment (type I I persistence) or are induced by treatment (type I persistence), as indicated by yellow question marks, and survive the duration of drug treatment. Once treatment is stopped, persisters re-establish a mixed population of sensitive and drug-tolerant cells, which remains susceptible to repeated drug exposure. (B) Shown are survival curves corresponding to populations in (A), plotted as the log number of cells over time during drug exposure.

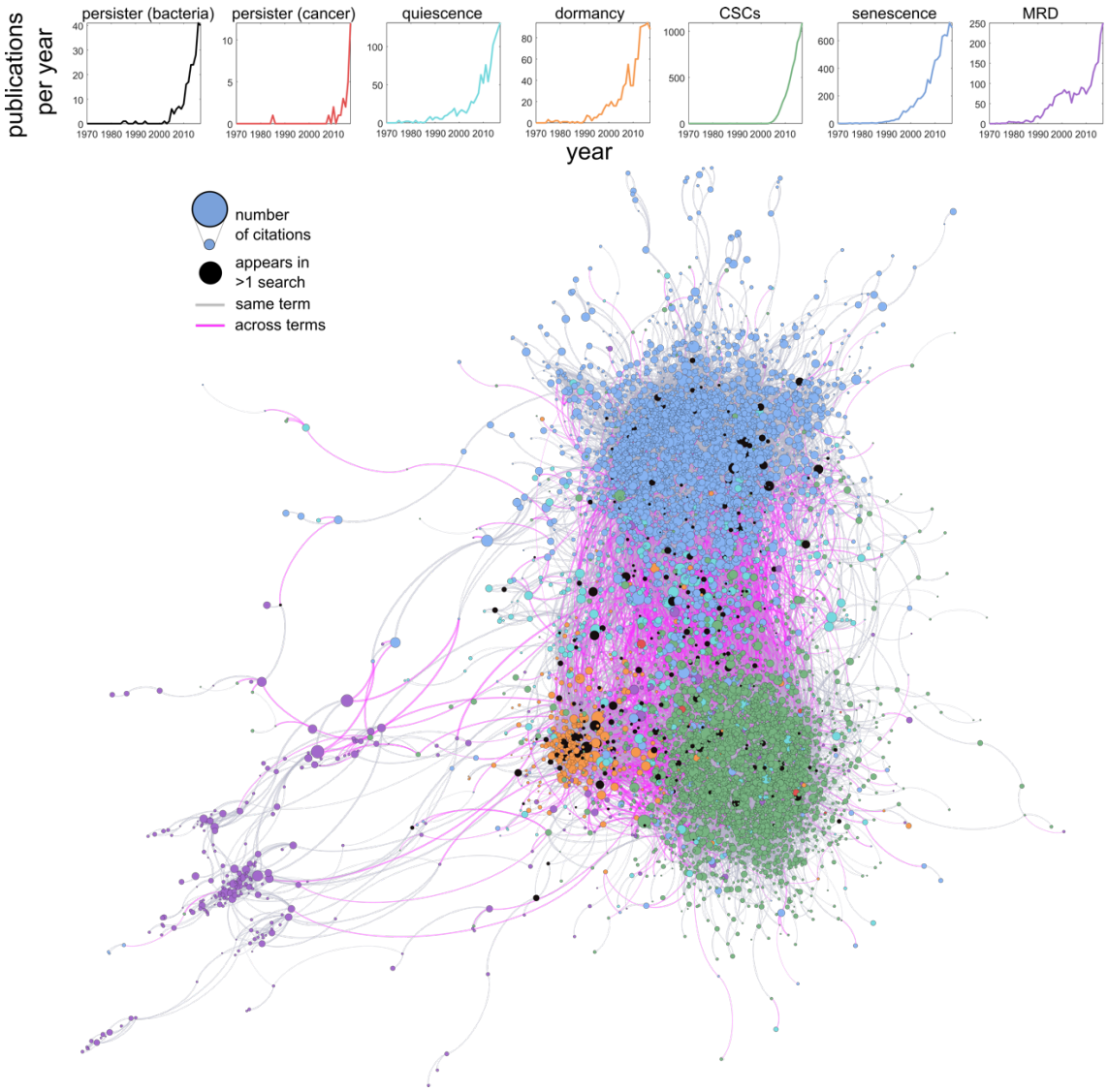

Figure 1.2: Diverse terminology and publication clusters in a literature citation network
(top) The number of publications per year is shown for six common terms used to describe similar phenomena related to drug tolerance (persistence, quiescence, dormancy, cancer stem cells (CSCs), senescence, and minimal residual disease (MRD)). (bottom) A cancer literature citation network is shown as a directed graph, where each node is a publication, the node's size represents its number of citations, the node's color represents its relevant term (colors same as top; black represents publications that contain >1 term), and edges are directed arcs (gray: within term, pink: across terms). Publications using the same term cluster together as shown in this force-directed layout (ForceAtlas 2).

Table 1.1: *in vitro* persister model systems in cancer
Published persister models in the literature, sorted by cancer origin and cell line name.

Cancer	Cell line	Target	Drug	Susceptibility	Reference
Breast	BT474	HER2	Lapatanib, Trastuzumab	BAD/BCL-XL	Moody *et al*.
Breast	BT474	HER2	Lapatanib, Carboplatin+Paclitaxel	GPX4	Hangauer *et al*.
Breast	EVSA-T	PI3K	PI3 kinase inhibitor	KDM5	Vinogradova *et al*.
Breast	SKBR3	HER2	Lapatanib	KDM5	Vinogradova *et al*.
Colon	Colo205	BRAF	Vemurafenib	KDM5	Vinogradova *et al*.
Gastric	GTL-16	MET	Crizotinib, Etoposide	ALDH1A1	Raha *et al*.
Gastric	MKN-4	MET	Crizotinib	ALDH1A1	Raha *et al*.
Lung	HCC827	EGFR	Erlotinib	BCL-2/BCL-XL, pSTAT3, SOX2	Fan *et al*., Rothenberg *et al*.
Lung	HCC827	EGFR	Gefitinib	OCT4, HIF1a, IGF1R	Kobayashi *et al*., Murakami *et al*.
Lung	PC9	EGFR	Erlotinib	BCL-2/BCL-XL, pSTAT3, KDM5, GPX4	Fan *et al*., Vinogradova *et al*., Hangauer *et al*.
Lung	PC9	EGFR	Gefitinib	IGF1R, KDM5, OCT4, HIF1a, IGF1R	Sharma *et al*., Kobayashi *et al*., Murakami *et al*.
Ovarian	JCRB		Carboplatin+Paclitaxel	GPX4	Hangauer *et al*.
Skin	A375	BRAF	Vemurafenib	GPX4	Hangauer *et al*.
Skin	Hs888	BRAF	AZ628	KDM5	Vinogradova *et al*.
Skin	M14	BRAF	AZ628	KDM5	Vinogradova *et al*.
T-ALL	DND-41		GSI (Compound E)	BRD4	Knoechel *et al*.
T-ALL	KOPT-K1		GSI (Compound E)	BRD4	Knoechel *et al*.

Methods

Citation network clustering and visualization

Publications for search terms (ex: "cancer AND persisters", "cancer AND quiescence") were saved from PubMed.gov as a list of PMIDs. To generate the list of publications which cited these works, each unique PMID was queried with rentrez (package version 1.2.2, feature: pubmed_pubmed_citedin) in R (version 3.6.0; R Core Team, 2019), and each citation was saved as an edge. For visualization purposes, publications with 2 or fewer citations and leaf vertices were dropped from further analysis. The graph was visualized using Gephi (version 0.9.2) with ForceAtlas 2 and Approximate Repulsion (Tolerance 1.0, Approximation 1.2, Scaling 2.0, Gravity 1.0).

References

Allison KR, Brynildsen M., Collins JJ (2011). Metabolite-enabled eradication of bacterial persisters by aminoglycosides. Nature, 473:216–220.

Amato SM, Brynildsen MP. (2015). Persister heterogeneity arising from a single metabolic stress. Current Biology, 25:2090–2098.

Amato SM, Fazen CH, Henry TC, Mok WWK, Orman MA, Sandvik EL, Volzing KG, Brynildsen MP. (2014). The role of metabolism in bacterial persistence. Frontiers in Microbiology, 5:1–9.

Amato SM, Orman MA, Brynildsen MP (2013). Metabolic Control of Persister Formation in Escherichia coli. Molecular Cell, 50:475–487.

Balaban NQ, Gerdes K, Lewis K, McKinney JD. (2013). A problem of persistence: Still more questions than answers? Nature Reviews Microbiology, 11:587–591.

Balaban NQ, Merrin J, Chait R, Kowalik L, Leibler S. (2004). Bacterial persistence as a phenotypic switch. Science, 305:1622.

Bigger JW. (1944). TREATMENT OF STAPHYLOCOCCAL INFECTIONS WITH PENICILLIN BY INTERMITTENT STERILISATION. Lancet, 244:497–500.

Borkowski O, Goelzer A, Schaffer M, Calabre M, Mäder U, Aymerich S, Jules M, Fromion V. (2016). Translation elicits a growth rate-dependent, genome-wide, differential protein production in Bacillus subtilis. Molecular Systems Biology, 12:870.

Brauner A, Fridman O, Gefen O, Balaban NQ. (2016). Distinguishing between resistance, tolerance and persistence to antibiotic treatment. Nature Reviews Microbiology, 14:320–330.

Chen J, Li Y, Yu TS, McKay RM, Burns DK, Kernie SG, Parada LF. (2012). A restricted cell population propagates glioblastoma growth after chemotherapy. Nature, 488:522–526.

Claudi B, Spröte P, Chirkova A, Personnic N, Zankl J, Schürmann N, Schmidt A, Bumann D. (2014). Phenotypic variation of salmonella in host tissues delays eradication by antimicrobial chemotherapy. Cell, 158:722–733.

Dixon SJ, Lemberg KM, Lamprecht MR, Skouta R, Zaitsev EM, Gleason CE, Patel DN, Bauer AJ, Cantley AM, Yang WS, *et al*. (2012). Ferroptosis: An iron-dependent form of nonapoptotic cell death. Cell, 149:1060–1072.

Fallahi-Sichani M, Becker V, Izar B, Baker GJ, Lin J, Boswell SA, Shah P, Rotem A, Garraway LA, Sorger PK. (2017). Adaptive resistance of melanoma cells to RAF inhibition via reversible induction of a slowly dividing de-differentiated state. Molecular Systems Biology, 13:905.

Fan W, Tang Z, Yin L, Morrison B, Hafez-Khayyata S, Fu P, Huang H, Bagai R, Jiang S, Kresak A, *et al*. (2011). MET-independent lung cancer cells evading EGFR kinase inhibitors are therapeutically susceptible to BH3 mimetic agents. Cancer Research, 71:4494–4505.

Fisher RA, Gollan B, Helaine S. (2017). Persistent bacterial infections and persister cells. Nature Reviews Microbiology, 15:453–464.

Flavahan WA, Drier Y, Liau BB, Gillespie SM, Venteicher AS, Stemmer-Rachamimov AO, Bradley E, Hospital MG, Chase C, Hospital MG. (2016). Insulator dysfunction and oncogene activation in IDH mutant gliomas. Nature, 529:110–114.

Fridman O, Goldberg A, Ronin I, Shoresh N, Balaban NQ. (2014). Optimization of lag time underlies antibiotic tolerance in evolved bacterial populations. Nature, 513:418–421.

Fung DKC, Chan EWC, Chin ML, Chan RCY. (2010). Delineation of a bacterial starvation stress response network which can mediate antibiotic tolerance development. Antimicrobial Agents and Chemotherapy, 54:1082–93.

Germain E, Castro-Roa D, Zenkin N, Gerdes K. (2013). Molecular Mechanism of Bacterial Persistence by HipA. Molecular Cell, 52:248–254.

Goldman A, Majumder B, Dhawan A, Ravi S, Goldman D, Kohandel M, Majumder PK, Sengupta S. (2015). Temporally sequenced anticancer drugs overcome adaptive resistance by targeting a vulnerable chemotherapy-induced phenotypic transition. Nature Communications, 6:1–13.

Gorrini C, Harris IS, Mak TW. (2013). Modulation of oxidative stress as an anticancer strategy. Nature Reviews Drug Discovery, 12:931–47.

Guler GD, Tindell CA, Pitti R, Wilson C, Nichols K, KaiWai Cheung T, Kim HJ, 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.

Gupta PB, Fillmore CM, Jiang G, Shapira SD, Tao K, *et al*. (2011). Stochastic state transitions give rise to phenotypic equilibrium in populations of cancer cells. Cell, 146:633–644.

Hangauer MJ, Viswanathan VS, Ryan MJ, Bole D, Eaton JK, Matov A, Galeas J, Dhruv HD, Berens ME, Schreiber SL, *et al*. (2017). Drug-tolerant persister cancer cells are vulnerable to GPX4 inhibition. Nature, 551:247-250.

Harms A, Maisonneuve E, Gerdes K. (2016). Mechanisms of bacterial persistence during stress and antibiotic exposure. Science, 354:aaf4268.

Hata AN, Niederst MJ, Archibald HL, Gomez-Caraballo M, Siddiqui FM, Mulvey HE, Maruvka YE, Ji F, Bhang HE, Krishnamurthy Radhakrishna V, *et al*. (2016). Tumor cells can follow distinct evolutionary paths to become resistant to epidermal growth factor receptor inhibition. Nature Medicine, 2016, 22:262–269.

Hui S, Silverman JM, Chen SS, Erickson DW, Basan M, Wang J, Hwa T, Williamson JR (2015). Quantitative proteomic analysis reveals a simple strategy of global resource allocation in bacteria. Molecular Systems Biology, 11:e784–e784.

Inde Z, Dixon SJ. (2017). The impact of non-genetic heterogeneity on cancer cell death. Critical Reviews in Biochemistry and Molecular Biology, 0:1–16.

Jordan NV, Bardia A, Wittner BS, Benes C, Ligorio M, Zheng Y, Yu M, Sundaresan TK, Licausi JA, Desai R, *et al*. (2016). HER2 expression identifies dynamic functional states within circulating breast cancer cells. Nature, 537:102–106.

Kaspy I, Rotem E, Weiss N, Ronin I, Balaban NQ, Glaser G. (2013). HipA-mediated antibiotic persistence via phosphorylation of glutamyl-tRNA-synthetase. Nature Communications, 4:1-7.

Knoechel B, Roderick JE, Williamson KE, Zhu J, Lohr JG, Cotton MJ, Gillespie SM, Fernandez D, Ku M, Wang H, *et al*. (2014). An epigenetic mechanism of resistance to targeted therapy in T cell acute lymphoblastic leukemia. Nature Genetics, 46:364–370.

Kobayashi I, Takahashi F, Nurwidya F, Nara T, Hashimoto M, Murakami A, Yagishita S, Tajima K, Hidayat M, Shimada N, *et al*. (2016). Oct4 plays a crucial role in the maintenance of gefitinib-resistant lung cancer stem cells. Biochem Biophys Res Commun, 473:125–132.

Kochanowski K, Gerosa L, Brunner SF, Christodoulou D, Nikolaev Y V, Sauer U. (2017). Few regulatory metabolites coordinate expression of central metabolic genes in Escherichia coli. Molecular Systems Biology, 13:903.

Kotte O, Zaugg JB, Heinemann M. (2010). Bacterial adaptation through distributed sensing of metabolic fluxes. Molecular Systems Biology, 6:355.

Kreso A, Dick JE. Evolution of the cancer stem cell model. Cell Stem Cell 2014, 14:275–291.

Kreso A, O'Brien CA, van Galen P, Gan OI, Notta F, Brown AMK, Ng K, Ma J, Wienholds E, Dunant C, *et al*. (2013). Variable clonal repopulation dynamics influence chemotherapy response in colorectal cancer. Science, 339:543–8.

Levin BR, Concepción-Acevedo J, Udekwu KI. (2014). Persistence: A copacetic and parsimonious hypothesis for the existence of non-inherited resistance to antibiotics. Current Opinion in Microbiology, 21:18–21.

Levin-Reisman I, Ronin I, Gefen O, Braniss I, Shoresh N, Balaban NQ (2017). Antibiotic tolerance facilitates the evolution of resistance. Science, 355:826–830.

Liau BB, Sievers C, Donohue LK, Gillespie SM, Flavahan WA, Miller TE, Venteicher AS, Hebert CH, Carey CD, Rodig SJ, *et al*. (2017). Adaptive Chromatin Remodeling Drives Glioblastoma Stem Cell Plasticity and Drug Tolerance. Cell Stem Cell, 20:233–246.e7.

Maisonneuve E, Castro-Camargo M, Gerdes K. (2013). (p)ppGpp controls bacterial persistence by stochastic induction of toxin-antitoxin activity. Cell, 154:1140–1150.

Mechler L, Herbig A, Paprotka K, Fraunholz M, Nieselt K, Bertram R. (2015). A novel point mutation promotes growth phase-dependent daptomycin tolerance in Staphylococcus aureus. Antimicrob Agents Chemother, 59:5366–5376.

Meylan S, Porter CBM, Yang JH, Belenky P, Gutierrez A, Lobritz MA, Park J, Kim SH, Moskowitz SM, Collins JJ. (2017). Carbon Sources Tune Antibiotic Susceptibility in Pseudomonas aeruginosa via Tricarboxylic Acid Cycle Control. Cell Chem Biol, 24:195–206.

Rothenberg SM, Concannon K, Cullen S, Boulay G, Turke AB, Faber AC, Lockerman EL, Rivera MN, Engelman JA, Maheswaran S, *et al*. (2015). Inhibition of mutant EGFR in lung cancer cells triggers SOX2-FOXO6 dependent survival pathways. eLife, 2015:1–25.

Mingay M, Chaturvedi A, Bilenky M, Cao Q, Jackson L, Hui T, Moksa M, Heravi-Moussavi A, Humphries RK, Heuser M, *et al*. (2017). Vitamin C-induced epigenomic remodelling in IDH1 mutant acute myeloid leukaemia. Leukemia, 32:11-20.

Moody SE, Schinzel AC, Singh S, Izzo F, Strickland MR, Luo L, Thomas SR, Boehm JS, Kim SY, Wang ZC, *et al*. (2014). PRKACA mediates resistance to HER2-targeted therapy in breast cancer cells and restores anti-apoptotic signaling. Oncogene, 34:2061–2071.

Mulcahy LR, Burns JL, Lory S, Lewis K. (2010). Emergence of *P. aeruginosa* strains producing high levels of persister cells in patients with cystic fibrosis. J Bacteriol, 192:6191–6199.

Murakami A, Takahashi F, Nurwidya F, Kobayashi I, Minakata K, Hashimoto M, Nara T, Kato M, Tajima K, Shimada N, *et al*. (2014). Hypoxia increases gefitinib-resistant lung cancer stem cells through the activation of insulin-like growth factor 1 receptor. PLoS One, 9:1–12.

Nguyen D, Joshi-Datar A, Lepine F, Bauerle E, Olakanmi O, Beer K, McKay G, Siehnel R, Schafhauser J, Wang Y, *et al*. (2011). Active starvation responses mediate antibiotic tolerance in biofilms and nutrient-limited bacteria. Science, 334:982–6.

Page R, Peti W. (2016). Toxin-antitoxin systems in bacterial growth arrest and persistence. Nature Chemical Biology, 12:208–214.

Pearl Mizrahi S, Gefen O, Simon I, Balaban NQ. (2016). Persistence to anti-cancer treatments in the stationary to proliferating transition. Cell Cycle, 15:3442–3453.

Pisco AO, Brock A, Zhou J, Moor A, Mojtahedi M, Jackson D, Huang S. (2013). Non-Darwinian dynamics in therapy-induced cancer drug resistance. Nature Communcications, 4:1–11.

Pisco AO, Huang S. (2015). Non-genetic cancer cell plasticity and therapy-induced stemness in tumour relapse: "What does not kill me strengthens me." British Journal of Cancer, 112:1725–1732.

Pu Y, Zhao Z, Li Y, Zou J, Ma Q, Zhao Y, Ke Y, Zhu Y, Chen H, Baker MAB, *et al*. (2016). Enhanced Efflux Activity Facilitates Drug Tolerance in Dormant Bacterial Cells. Mol Cell, 62:284–294.

Radzikowski JL, Schramke H, Heinemann M. (2017). Bacterial persistence from a system-level perspective. Curr Opin Biotechnol, 46:98–105.

Radzikowski JL, Vedelaar S, Siegel D, Ortega ÁD, Schmidt A, Heinemann M. (2016). Bacterial persistence is an active σS stress response to metabolic flux limitation. Mol Syst Biol, 12:882.

Raha D, Wilson TR, Peng J, Peterson D, Yue P, Evangelista M, Wilson C, Merchant M, Settleman J. (2014). The Cancer Stem Cell Marker Aldehyde Dehydrogenase Is Required to Maintain a Drug-Tolerant Tumor Cell Subpopulation. Cancer Res, 74:3579–3590.

Ramirez M, Rajaram S, Steininger RJ, Osipchuk D, Roth MA, Morinishi LS, Evans L, Ji W, Hsu C-H, Thurley K, *et al*. (2016). Diverse drug-resistance mechanisms can emerge from drug-tolerant cancer persister cells. Nature Communications, 7:10690.

Roesch A, Fukunaga-Kalabis M, Schmidt EC, Zabierowski SE, Brafford PA, Vultur A, Basu D, Gimotty P, Vogt T, Herlyn M. (2010). A Temporarily Distinct Subpopulation of Slow-Cycling Melanoma Cells Is Required for Continuous Tumor Growth. Cell, 141:583–594.

Roesch A, Vultur A, Bogeski I, Wang H, Zimmermann KM, *et al*. (2013). Overcoming intrinsic multidrug resistance in melanoma by blocking the mitochondrial respiratory chain of slow-cycling JARID1B(high) cells. Cancer Cell, 23:811–25.

Sabharwal SS, Schumacker PT. (2014). Mitochondrial ROS in cancer: initiators, amplifiers or an Achilles' heel? Nat Rev Cancer, 14:709–721.

Schmidt A, Kochanowski K, Vedelaar S, Ahrné E, Volkmer B, Callipo L, Knoops K, Bauer M, Aebersold R, Heinemann M. (2016). The quantitative and condition-dependent Escherichia coli proteome. Nat Biotechnol, 34:104–10.

Sciacovelli M, Gonçalves E, Johnson TI, Zecchini VR, da Costa ASH, Gaude E, Drubbel AV, Theobald SJ, Abbo SR, Tran MGB, *et al*. (2016). Fumarate is an epigenetic modifier that elicits epithelial-to-mesenchymal transition. Nature, 537:544–547.

Scott M, Gunderson CW, Mateescu EM, Zhang Z, Hwa T. (2010). Interdependence of Cell Growth and Gene Expression: Origins and Consequences. Science, 330:1099–1102.

Scott M, Hwa T. (2011). Bacterial growth laws and their applications. Curr Op Biotech, 22(4):559-65.

Shaffer SM, Dunagin MC, Torborg SR, Torre EA, Emert B, Krepler C, Beqiri M, Sproesser K, Brafford PA, Xiao M, *et al*. (2017). Rare cell variability and drug-induced reprogramming as a mode of cancer drug resistance. Nature, 546:431–435.

Shan Y, Gandt AB, Rowe SE, Deisinger JP, Conlon BP, Lewis KIM, Brown Gandt A, Rowe SE, Deisinger JP, Conlon BP, *et al*. (2017). ATP-Dependent Persister Formation in Escherichia coli. mBio, 8:e02267-16.

Sharma SV, Lee DY, Li B, Quinlan MP, Takahashi F, Maheswaran S, McDermott U, Azizian N, Zou L, Fischbach MA, *et al*. (2010). A Chromatin-Mediated Reversible Drug-Tolerant State in Cancer Cell Subpopulations. Cell, 141:69–80.

Su Y, Wei W, Robert L, Xue M, Tsoi J, Garcia-Diaz A, Homet Moreno B, Kim J, Ng RH, Lee JW, *et al*. (2017). Single-cell analysis resolves the cell state transition and signaling dynamics associated with melanoma drug-induced resistance. Proc Natl Acad Sci, 114(52):13679-13684.

Torre LA, Bray F, Siegel RL, Ferlay J, Lortet-Tieulent J, Jemal A. (2015). Global cancer statistics, 2012. CA Cancer J Clin, 65:87–108.

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

Van Den Bergh B, Michiels JE, Wenseleers T, Windels EM, Boer P Vanden, Kestemont D, De Meester L, Verstrepen KJ, Verstraeten N, Fauvart M, *et al*. (2016). Frequency of antibiotic application drives rapid evolutionary adaptation of Escherichia coli persistence. Nat Microbiol, 1:1–7.

Verstraeten N, Knapen WJ, Kint CI, Liebens V, Van den Bergh B, Dewachter L, Michiels JE, Fu Q, David CC, Fierro AC, *et al*. (2015). Obg and Membrane Depolarization Are Part of a Microbial Bet-Hedging Strategy that Leads to Antibiotic Tolerance. Mol Cell, 59:9–21.

Viale A, Draetta GF. (2016). Metabolic features of cancer treatment resistance. Recent Results in Cancer Research, 207:135–156.

Viale A, Pettazzoni P, Lyssiotis CA, Ying H, Sánchez N, Marchesini M, Carugo A, Green T, Seth S, Giuliani V, *et al*.: Oncogene ablation-resistant pancreatic cancer cells depend on mitochondrial function. Nature 2014, 514:628–632.

Vinogradova M, Gehling VS, Gustafson A, Arora S, Tindell CA, Wilson C, Williamson KE, Guler GD, Gangurde P, Manieri W, *et al*. (2016). An inhibitor of KDM5 demethylases reduces survival of drug-tolerant cancer cells. Nat Chem Biol, 12:531–8.

Viswanathan VS, Ryan MJ, Dhruv HD, Gill S, Eichhoff OM, Seashore-Ludlow B, Kaffenberger SD, Eaton JK, Shimada K, Aguirre AJ, *et al*. (2017). Dependency of a therapy-resistant state of cancer cells on a lipid peroxidase pathway. Nature, 547:453–457.

Wilson TR, Fridlyand J, Yan Y, Penuel E, Burton L, Chan E, Peng J, Lin E, Wang Y, Sosman J, *et al*. (2012). Widespread potential for growth-factor-driven resistance to anticancer kinase inhibitors. Nature, 487:505–509.

Wolf DA. (2014). Is Reliance on Mitochondrial Respiration a “Chink in the Armor” of Therapy-Resistant Cancer? Cancer Cell, 26:788–795.

Chapter 2

Loss of TET2 affects proliferation and drug sensitivity through altered dynamics of cell-state transitions

Main contributing authors

Leanna Morinishi[1], Karl Kochanowski[2], Ross L. Levine[3,4,5], Lani F. Wu[2] & Steven J. Altschuler[2]

Affiliations

[1] Bioinformatics Graduate Group, University of California, San Francisco

[2] Department of Pharmaceutical Chemistry, University of California, San Francisco

[3] Human Oncology and Pathogenesis Program and Center for Hematologic Malignancies, Memorial Sloan Kettering Cancer Center

[4] Department of Medicine, Leukemia Service, Memorial Sloan Kettering Cancer Center

[5] Center for Epigenetics Research, Memorial Sloan Kettering Cancer Center

Abstract

A persistent puzzle in cancer biology is how mutations, which neither alter canonical growth signaling pathways nor directly interfere with drug mechanism, can still recur and persist in tumors. One notable example is the loss-of-function mutation of the DNA demethylase TET2 in acute myeloid leukemias (AMLs) that frequently persists from diagnosis through remission and relapse (Rothenberg-Thurley *et al*., 2018; Corces-Zimmerman *et al*., 2014; Nibourel *et al*., 2010), but whose fitness advantage in the setting of anti-leukemic chemotherapy is unclear. Here we use paired isogenic human AML cell lines to show that TET2 loss-of-function alters the dynamics of transitions between differentiated and stem-like states. A conceptual mathematical model and experimental validation suggest these altered cell-state dynamics can benefit the cell population by slowing population decay during drug treatment and lowering the number of survivor cells needed to re-establish the initial population. These studies shed light on the functional and phenotypic effects of a TET2 loss-of-function in AML, illustrate how a single gene mutation can alter a cells' phenotypic plasticity, and open up new avenues in the development of strategies to combat AML relapse.

Introduction

A major challenge in cancer biology is to understand the function of recurrent mutations in the emergence of tumors or response to drug therapy. Some mutations clearly benefit cancer populations either by: altering regulation of growth signaling or programmed cell death (e.g. mutations in p53 or TGFβ signaling, Sanchez-Vega *et al*., 2018), or by directly interfering with drug effect (e.g. acquired EGFR T790M resistance mutations in response to EGFR tyrosine kinase inhibitor therapy, Ma, Wei and Song, 2011). However, for many observed mutations it is unclear how they affect either proliferation or drug resistance.

One example are mutations in the DNA demethylase TET2, found in ~15-20% of de novo AMLs (i.e., AMLs in patients with no clinical history of myelodysplastic syndrome (MDS) and no prior exposure to potentially leukemogenic therapies; Cheson *et al*., 2003; Nibourel *et al*., 2010; Metzeler *et al*., 2011; Cancer Genome Atlas Research Network, 2013; Moran-Crusio *et al*., 2011). Mutated TET2 is associated with both pre-leukemic states, such as clonal hematopoiesis (a condition where a clone becomes overrepresented in the blood; Busque *et al*., 2012, Corces-Zimmerman *et al*., 2014, Potter *et al*., 2019), and mutational persistence and adverse outcome in human AML (Rothenberg-Thurley *et al*., 2018; Ding *et al*., 2011). However, due to its ability to alter DNA methylation genome-wide, the mechanisms by which TET2 mutation confers a benefit to AML cancer cell populations remain unclear. One possibility is that epigenetic variation in cancer cells increases phenotypic flexibility, enabling tumor evolution and progression (Feinberg *et al*., 2016; Flavahan *et al*., 2017). Indeed, recent studies have shown that loss of epigenetic effectors such as KDM5A in breast cancer (Hinohara *et al*., 2018) can affect drug sensitivity by enabling greater cell heterogeneity. Here we investigate whether TET2 loss similarly affects cell fitness in AML using an integrated approach of mathematical modeling and experimentation in paired WT and TET2-mutant isogenic human AML cell lines. We discover that TET2 mutation alters the dynamics of transitions between distinct stem-like and differentiated cell states, which enhances population fitness in chemotherapy and lowers the number of cells needed to establish a cell population.

Results

TET2 loss-of-function mutation renders AML cell populations more stem-cell like

To investigate the consequences of TET2 mutation, we chose to compare two pairs of isogenic human myeloblast cell lines, each expressing wildtype or mutant TET2 (Figure 2.1). The AML cell lines KG1 and Thp1 (Kunimoto *et al*., 2018) were selected as they express wildtype TET2 ($TET2^{WT}$) but do not express mutant FLT3, which is known to have synergistic epigenetic effects (Shish *et al*., 2015). In human AMLs, TET2 is often mutated in AML with truncating mutations or missense mutations in its catalytic domain (Hirsch *et al*., 2018), resulting in TET2 loss-of-function (Smith *et al*., 2010). Therefore, isogenic cell lines were created by knocking out TET2 in the chosen cell lines ($TET2^{KO}$, Methods). Loss of wildtype TET2 expression was confirmed *via* RT-qPCR of the TET2 transcript and immunoblotting of the N-terminus of the protein (Figure A.1). To confirm that loss of TET2 has the expected effect on DNA methylation (Yamazaki *et al*., 2015; Asmar *et al*., 2013; Rasmussen *et al*., 2015), we performed DNA methylation profiling. As expected, $TET2^{KO}$ cell lines display a significantly higher degree of overall hypermethylation (t-test, p-values: KG1 <2.2e-16 and Thp1 2.9e-4) compared to their WT counterparts, with high reproducibility across replicates (Figure 2.1, Figure A.2, Figure A.3).

To gain an unbiased overview of the molecular changes induced by $TET2^{KO}$, we examined the transcriptomic and epigenomic profiles of TET2 WT and KO cell populations (see Appendix A). RNAseq analysis identified ~300 gene transcripts that are similarly differentially expressed in both $TET2^{KO}$ cell lines (fold-change>2, Figure A.2). Notably, reduced expression of myeloid differentiation markers (ITGAM, CORO1A, Wald test, BH adjusted p-values 9.99e-9 and 7.73e-2, Figure 2.1, Figure A.2) and increased expression of markers associated with leukemic progenitor cells ($CD38^-$, $HLA\text{-}DRA^-$, TAL1, Wald test, BH adjusted p-values 3.83e-26, 1.74e-16, and 3.90e-7, Figure 2.1, Figure A.2; Chan *et al*., 2012; Kanehisa 2000; Kanehisa *et al*., 2019; Nishioka *et al*., 2013; Vagapova *et al*., 2018) were observed. These data are consistent with analysis of TCGA LAML (phs000178.v1.p1) and OHSU (Tyner *et al*., 2018) datasets, which showed that TET2 expression is strongly correlated with genes expressed in the granulocytic lineage (Fisher's exact test, adjusted p-values 8e-25 and 3.7e-19, Tables A.1-A.2, Methods).

Differentially expressed genes in TET2^{KO} cells were found to be highly enriched for targets of Runx1, a hematopoietic regulator known to promote stemness and myeloid fate decisions (Fisher's exact test, p-value 7.76e-6, Figure A.2, Tables A.3-A.4) (Kuleshov *et al*., 2016; Ran *et al*., 2013). Consistently, analysis of differentially methylated regions showed the proximal promoter of Runx1 to be significantly affected in TET2^{KO} (Wald test, minimum adjusted p-value 1.22e-27, Figure A.2), with a concomitant increase in the expression the Runx1a isoform (Figure A.2). Methylcellulose colony forming unit assays of TET2^{KO} cells also show increased numbers of colonies associated with oligopotent progenitor cells compared to TET2^{WT} controls (CFUGEMMs, Figure 2.1, Figure A.4). Overall, these data suggested that TET2^{KO} populations acquire more stem-like signatures, which was further validated by comparing their DNA methylation profiles to those of normal hematopoietic progenitors and leukemic stem cells (LSCs, Figure A.4; Horvath, 2013; Jung *et al*., 2015).

The TET2^{KO}-mediated acquisition of stem-like signatures was much more pronounced in KG1 than Thp1 in terms of expression profile and potential to form diverse myeloid lineages in colony-forming assays (Figure A.4). This is presumably because the monocyte-like cells, Thp1, are already more differentiated than the myeloblast-like cells, KG1, regardless of TET2 mutational status. We therefore chose to focus on KG1 cells for the remainder of this work.

TET2 mutation changes dynamics of cell-state switching

To test whether this change in stem-like molecular signatures is caused by an increase in the fraction of cells with CD34^{hi}CD38^{lo} surface marker expression (a classic LSC-like profile; Costello et al, 2000; Gerber *et al*., 2012; Nishioka *et al*., 2013; Zeijlemaker *et al*., 2018) in TET2^{KO} compared to TET2^{WT} cells, we measured CD34 and CD38 surface marker expression in the population using flow cytometry. Quantification of KG1 CD34/CD38 expression revealed that the fraction of CD34^{hi}CD38^{lo} cells was indeed increased in TET2^{KO} cell populations (Figure 2.2, Figure A.5).

To confirm that KG1 CD34^{hi}CD38^{lo} cells are more stem-like than CD34^{hi}CD38^{hi}, we measured the expression of genes associated with hematopoietic differentiation. As observed from bulk RNAseq (Fig

1C), TET2^{WT} cells compared to TET2^{KO} show a general shift towards higher levels of expression for pro-differentiation markers. Consistently, after fractionating cells by CD38 expression we noted that differentiation markers again increased with increasing levels of CD38 expression across both cell lines (trend seen for ITGAM, CORO1A, and TAL1, though not for HOXA5; Figure A.6). Further, growth rate decreased with increasing levels of CD38 expression (Figure A.7) across KG1 TET2^{KO} and TET2^{WT} cells. Together, these data further support that CD34^{hi}CD38^{lo} cells are more stem-like and self-renewing, whereas CD34^{hi}CD38^{hi} cells are more differentiated and non-dividing.

How does TET2^{KO} alter the ratio between stem-like and differentiated cells? One possibility is that TET2^{KO} simply reduces the rate of differentiation. To test this, subpopulations of stem-like (L, CD34^{hi}CD38^{lo}) and differentiated (H, CD34^{hi}CD38^{hi}) cells were sorted in both TET2^{WT} and TET2^{KO} populations, and the ratio of L and H cells was monitored over time (Figure 2.2, Figure A.8) and to steady-state (Figure A.9). Indeed, sorted stem-like cell populations repopulated the differentiated state more slowly in TET2^{KO} (Figure 2.2, Figure A.8). Surprisingly, in TET2^{KO}, the sorted differentiated cell populations repopulated the stem-cell like state more rapidly than TET2^{WT} populations (Figure 2.2, Figure A.8). These data suggest that TET2^{KO} facilitates the reversible switching from differentiated to stem-like states in AML cells and provides intuition for why TET2^{KO} has a larger fraction of stem-like cells in the population.

Mathematical modeling illustrates consequences of altered cell-state dynamics for survival of cell population

What are the functional consequences of having such altered cell-state dynamics, and in particular a higher fraction of stem-like cells? To address this question, we developed a simple, conceptual mathematical model with two cell states (Figure 2.3, second panel): a stem-cell-like state L, and a more differentiated state H. This model is fully characterized by a linear, homogeneous system of ordinary differential equations (ODEs) with 6 parameters with values ≥ 0 (Methods). Parameters are denoted by:

d_L and d_H the death rates of the L and H states, respectively; g_L and g_H, the rates at which the L and H states proliferate; and r_{LH} and r_{HL} the transition rates from L to H and H to L, respectively.

Using this model, we asked under which circumstances the observed altered cell-state dynamics – specifically an increased switch rate towards a stem-like state – would benefit a cell population. First, we focused on the impact of such altered cell-state dynamics during drug treatment (i.e. high death rates). In this case, the time to population collapse is dominated by the largest negative eigenvalue; the less negative the eigenvalue, the slower the population collapse. By computing the eigenvalues for the ODE, it can be seen that either increasing r_{HL} and/or decreasing r_{LH} slows population decay and "benefits" a drug-treated cancer population as long as (Figure 2.3, first panel; Methods):

$$(g_L - d_L) > (g_H - d_H) \quad \textbf{(Inequality 1)}$$

This inequality simply compares net production rates (proliferation rate minus death rate) and requires it to be higher at L than H. In fact, while Inequality 1 is true, increasing r_{HL}, or decreasing r_{LH} will eventually slow population decay (arrows in phase plane diagram below; see Methods for details). Thus, as long as the net production rate of the stem-like state L exceeds that of the differentiated state H, altered cell state dynamics (i.e. towards a higher fraction of stem-like cells) will always benefit the cell population. This finding is in line with recent studies suggesting the possibility that cancer cells can reversibly transit between stem and differentiated states with different drug sensitivities (Jordan *et al*., 2016; Su *et al*., 2017; Gupta *et al*., 2011).

Importantly, the same inequality also captures other cases in which switching to a stem-cell-like state is either beneficial or detrimental (Figure 2.3, third and fourth panels). If cells rarely die, such as in a non-drug-treated condition ($d_L \approx 0,\ d_H \approx 0$) (**Case 1**), inequality 1 simplifies to

$$g_L > g_H$$

Here, switching to the L state (increasing r_{HL}) is beneficial for the cell population when the proliferation rate g_L of L is higher than the proliferation rate g_H of H, which is likely given the lower propensity of differentiated cells to divide. In this case, we expect cells carrying mutations that increase r_{HL}

to take over the population. Conversely, if the differentiated cell state H were protected from the adverse effects of drug treatment ($d_H \approx 0$) (**Case 2**), inequality 1 simplifies to

$$g_L - d_L > g_H$$

Here, increasing r_{HL} is only beneficial for the cell population if the net production rate of L is larger than the proliferation rate g_H of the differentiated state H, which is unlikely for high doses of drug ($d_L \gg 0$). In this case, we expect cells carrying mutations that increase r_{HL} to be depleted from the population over time.

Overall, this conceptual mathematical model suggests that altered cell-state dynamics – specifically an increased switch rate towards a stem-like state – can indeed benefit the cell population (both in presence or absence of drug treatment), as long as it falls within a parameter regime outlined by inequality 1 (i.e. net production rate of stem-like state exceeds that of the differentiated state).

Experimental validation of modeling predictions

Does the TET2 mutation put the AML cancer population in this predicted advantageous parameter regime? Since the proliferation, death and switching rates cannot be disentangled directly from measurements, we estimated these rates by fitting our model to time course data of sorted $TET2^{KO}$ and WT populations in the presence of cytosine arabinoside (AraC or Cytarabine, a common first-line chemotherapy drug for AML; Figure 2.4, Figure A.8, see Methods). These parameter estimates confirmed that the proliferation rate during drug treatment of stem-like cells far exceeds that of differentiated cells, both for $TET2^{KO}$ and WT populations (Figure 2.4). Thus, both populations are in a regime where the net production of stem-like cells outstrips the net production of differentiated cells in drug treatment (as required by Inequality 1), and therefore switching to the stem-like state is more advantageous. The $TET2^{KO}$ populations have a higher $H \rightarrow L$ switching rate and a lower $L \rightarrow H$ switching rate; hence, the model predicts that the $TET2^{KO}$ population, as a whole, is more drug resistant. Indeed, drug treatment experiments confirmed that $TET2^{KO}$ populations are less sensitive to AraC and doxorubicin than $TET2^{WT}$ populations (Figure 2.4,

Figure A.8). We also used the same experimental setup to determine model parameters for different drug conditions and different CD38 subpopulations (Figure A.8, Figure A.9). Our key results – Inequality 1 holds, TET2^{KO} cells have a higher transit rate r_{HL} from H to L, and TET2^{WT} cells have a higher transit rate r_{LH} from L to H – held true across experiments (Figure A.11).

Further analysis of the model suggested two additional predictions. First, the model suggested that a population with a higher fraction of differentiated cell states will show reduced population survival in drug treatment (Figure 2.4). To test this conjecture, two effectors known to enrich AML cell populations for the differentiated state (without altering cell death, Figure A.12) were used, namely the inflammatory stimulus interferon gamma (IFNg) and the aldehyde dehydrogenase inhibitor disulfiram (DS) (Amici *et al*., 2018; Xu *et al*., 2017). We confirmed that treatment with IFNg or DS enriched both TET2^{KO} and WT populations for the differentiated cell state after 72 hours of exposure (Methods, Figure 2.4, Figure A.12). Moreover, both effectors increased the efficacy of AraC treatment in both TET2^{KO} and WT populations, while several effectors not known to affect cell-state transitions did not alter AraC sensitivity (Figure 2.4, Figure A.12, Table A.5).

Second, the model predicted that the increased ability of TET2^{KO} to revert to a stem-like cell state (with its higher proliferation potential), will allow a TET2^{KO} population to better regrow out of drug than a TET2^{WT} population (Figure 2.4). To test this prediction, we compared the number of colonies formed by TET2^{KO} and TET2^{WT} populations in methylcellulose assays, which assesses the ability of isolated cells to reform a colony (Figure 2.4). As predicted, sorted stem-like cell populations showed increased cell colony numbers in both TET2^{KO} and TET2^{WT} populations. Moreover, unsorted TET2^{KO} populations formed approximately 2x more colonies than their TET2 WT counterparts (Figure 2.4), highlighting their increased potential for population renewal. Thus, our results suggest that TET2^{KO} populations have a higher likelihood to regrow an AML population from few surviving cells in the setting of therapeutic perturbations.

Discussion

Taken together, our results reveal that mutation of the epigenetic modifier TET2 in AML cell populations alters the transition dynamics between stem-like and differentiated cell states. These altered cell-state dynamics confer several benefits to the population. First, TET2 mutation enables differentiated cells to switch to a stem-like state, which has a higher net production rate either in or out of drug, thus increasing population survival. Second, TET2 mutation increases the propensity to regrow an AML population after drug treatment from few surviving cells, due to the increased ability of $TET2^{KO}$ cells to revert back to a proliferative stem-like cell state (Figure 2.4). These findings provide a rationale for why the detection of TET2 mutants in patients in remission is strongly associated with a higher chance of relapse (Rothenberg-Thurley *et al*., 2018; Ding *et al*., 2012).

The key results from this study of a human $CD34^{+}CD38^{+}$ myeloid cell line (KG1) are consistent with past murine studies showing that TET2 mutation alters hematopoietic stem cell self-renewal and differentiation in bulk (Moran-Crusio *et al*., 2011). Future studies will be necessary to determine the generalizability of the cell state-switching phenomenon in other hematopoietic cell types and *in vivo*. Additionally, our study was limited to cell states defined by well-known cell surface markers. Incorporating powerful, unbiased approaches such as scRNAseq and analyses measuring RNA velocity (La Manno *et al*., 2018) could potentially uncover novel cell states and state switching dynamics that are relevant for drug survival.

An interesting question raised by this study is the degree to which the state switching rates are cell intrinsic or extrinsically determined. In the process of confirming our model, we indeed found that the transition of $TET2^{KO}$ cells from $CD38^{hi}$ to $CD38^{lo}$ could be altered by effectors like DS and IFNg (Figure 2.4). To characterize the extent to which cell-state switching rates are affected by signaling between the cells themselves, we grew KG1 $TET2^{WT}$ cells in media conditioned by $TET2^{KO}$ cells and *vice versa*. The distribution of CD38 expression in $TET2^{WT}$ cells was not affected by $TET2^{KO}$ conditioned media, but interestingly CD38 expression clearly increased for $TET2^{KO}$ cells grown in $TET2^{WT}$ conditioned media (Figure A.13) with decreased switching rates from $CD38^{hi}$ to $CD38^{lo}$ states (Figure A.13). Together, these

results suggest that the switching rates of $TET2^{KO}$ cells can be influenced by external perturbations, while switching rates of $TET2^{WT}$ are more stable. Future studies will be needed to identify key soluble factors that distinguish $TET2^{KO}$ conditioned media from $TET2^{WT}$ and, more generally, to fully test the extent to which cell state switching rates are intrinsically determined.

This work provides further evidence that mutations which modulate switching rates between more and less drug-resistant states may provide an "evolutionary shortcut" to counteract the adverse effect of drug treatment (Jordan *et al*., 2016; Su *et al*., 2017; Gupta *et al*., 2011). The mechanisms driving this dynamic tumor heterogeneity will require deeper study, however, therapeutic strategies incorporating this knowledge may better counteract tumor evolution in response to treatment. For example, a natural extension of the work presented here would be to use the model and understanding of cancer cell state switching dynamics to determine ideal dosing strategies (dosage, time course, pulsed administration, and so on) to help therapeutically manage patient outcomes. A recently published review made similar suggestions for future therapeutic development in melanoma (Bai *et al*., 2019).

Our results have implications for AML mutations beyond TET2: the conceptual mathematical model shows that as long as stem-like and differentiated cell states differ in their net cell production during drug treatment (i.e., Inequality 1 is fulfilled), any other mutation increasing the transition rate towards the stem-like state will be beneficial as well. Thus, we conjecture that mutations in other epigenetic modifiers might confer a fitness advantage to drug-treated AML populations through similar mechanisms. In cancer, the target space for mutations that directly interfere with drug function is likely smaller than the target space for dysregulation of cell states; this is known to be the case in bacteria, where the number of genes conferring increased antibiotic tolerance far exceeds the number of genes conferring drug resistance by interfering directly with drug mechanism (Girgis, Harris, and Tavazoie, 2012; Brauner *et al*., 2016). Future studies into characterizing cell states and transition dynamics will provide new insight into why certain mutations are selected for, and persist in, cancer, and led to new therapeutic approaches which increase the efficacy of current cancer therapies.

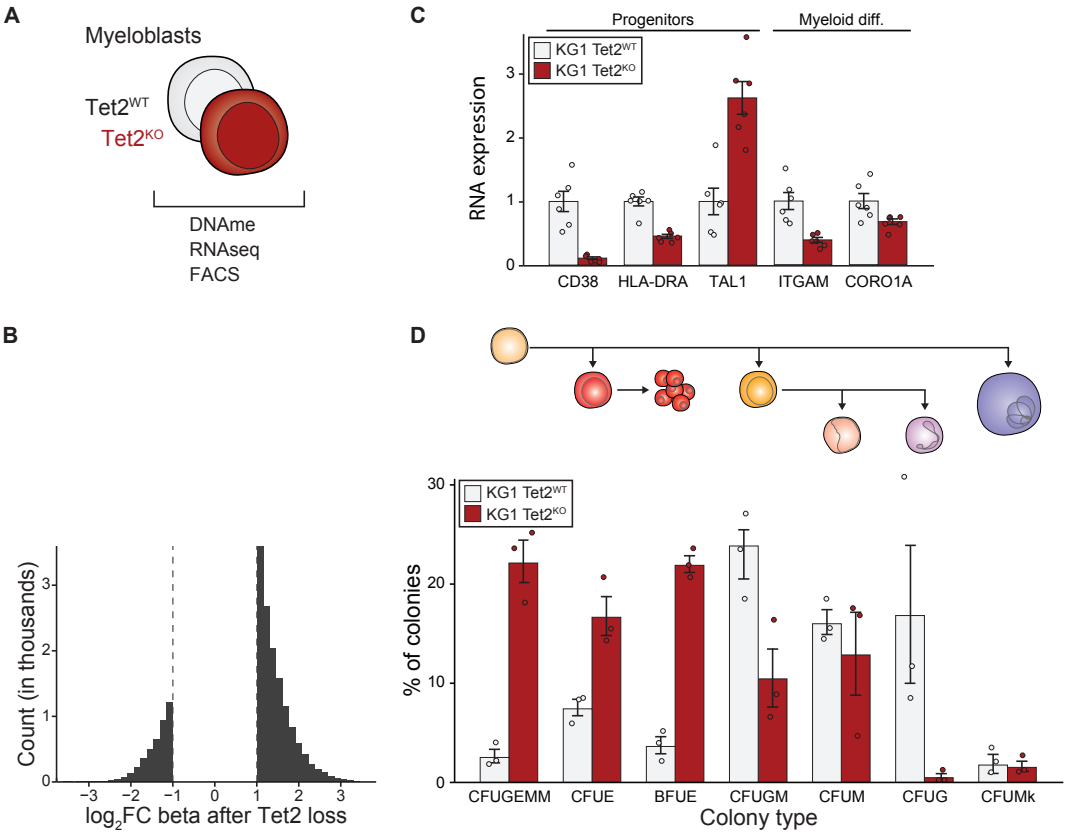

Figure 2.1: TET2^{KO} cells are more stem-like than TET2^{WT} isogenic counterparts
(A) Overview of molecular profiling performed on isogenic TET2 mutant AML cell lines. (B) TET2^{KO} cell lines are more hypermethylated than their wildtype counterparts. Shown is a histogram of log_2fold-change (CpG beta values) in TET2^{KO} cell lines for differences (relative to parental cell lines) greater than 2-fold from a paired analysis. (C) Genes that modulate differentiation and stemness as measured by RNAseq are differentially expressed after TET2^{KO} in both KG1 and Thp1 (Figure A.2, mean ± s.e., n=6). TET2^{KO} cell lines show decreased expression of myeloid commitment markers (ITGAM, CORO1A) and increased expression of markers associated with stemness (CD38$^-$, HLA-DRA$^-$, TAL1) compared to TET2^{WT}. (D) KG1 and Thp1 (Figure A.4) TET2^{KO} cell lines produce a higher percentage of colonies associated with oligopotent progenitor cells (CFUGEMM) compared to TET2^{WT} cells (mean ± s.e., n=3) in methylcellulose colony forming assays.

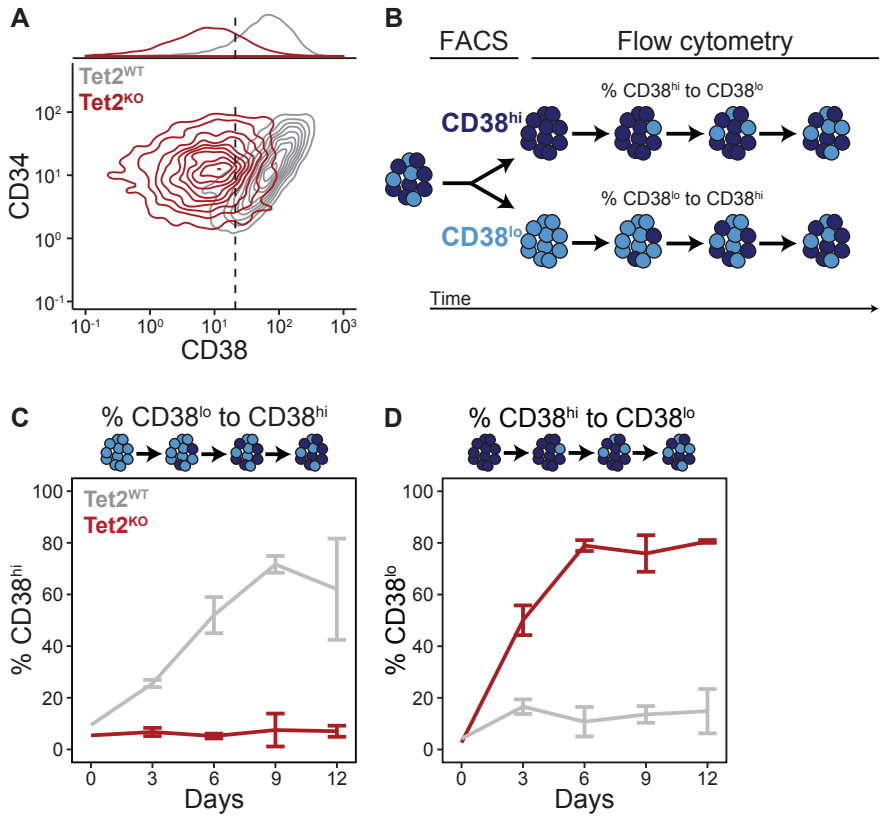

Figure 2.2: TET2^{KO} increases the propensity of differentiated (CD34^{hi}CD38^{hi}) cells to switch to a more stem-like (CD34^{hi}CD38^{lo}) cell state

(A) KG1 TET2^{KO} cell lines show a shift in CD38 surface marker expression. The threshold for calling cells as CD38^{lo} or CD38^{hi} (vertical dashed line) is defined by negative controls, such that 99.5% of unstained cells are classified as CD38^{lo} (Figure A.5). (B) Schematic illustration of phenotypic segregation and flow cytometry experiment to quantify CD38 expression over time. Cells from either KG1 TET2^{WT} or TET2^{KO} cell lines are sorted by CD38 expression via FACS (fluorescence-activated cell sorting) into CD38^{lo} or CD38^{hi} based on the threshold in (A), and changes in CD38 expression are assessed every 3 days by flow cytometry (Figure A.8). (C-D) The percent CD38^{hi} cells in originally pure populations of sorted CD38^{lo} cells (C) or the percent CD38^{lo} cells in originally pure populations of sorted CD38^{hi} cells (D) for both KG1 TET2^{WT} or TET2^{KO} cells after 0, 3, 6, 9, and 12 days of growth in drug-free media (see Figure A.8; mean $\pm$ s.e., n=2).

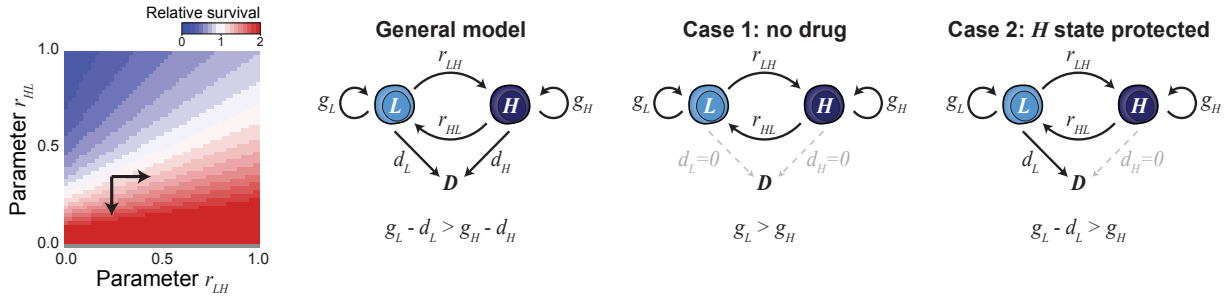

Figure 2.3: Mathematical model reveals advantageous and disadvantageous parameter regimes for cell-state switching

(left) If model parameters satisfy Inequality 1, both increasing r_{HL} and/or decreasing r_{LH} (black arrows) slows population decay and benefits a drug-treated cancer population. (General model) Schematic representation of mathematical model with three cell states: a stem-cell-like state L, a differentiated state H, and an irreversible cell death state. Parameters d_L and d_H represent the death rates of the L and H states; g_L and g_H the proliferation rates; and r_{LH} and r_{HL} the transition rates from L to H and H to L. (Case 1) If cells rarely die ($d_L \approx 0$, $d_H \approx 0$) and the proliferation rate g_L is higher than g_H, increasing r_{HL} is always beneficial. (Case 2) Conversely, if one state were protected from drug effect ($d_H \approx 0$), increasing r_{HL} is only beneficial in the unlikely scenario when net production rate of the other state L is larger than its proliferation rate g_H.

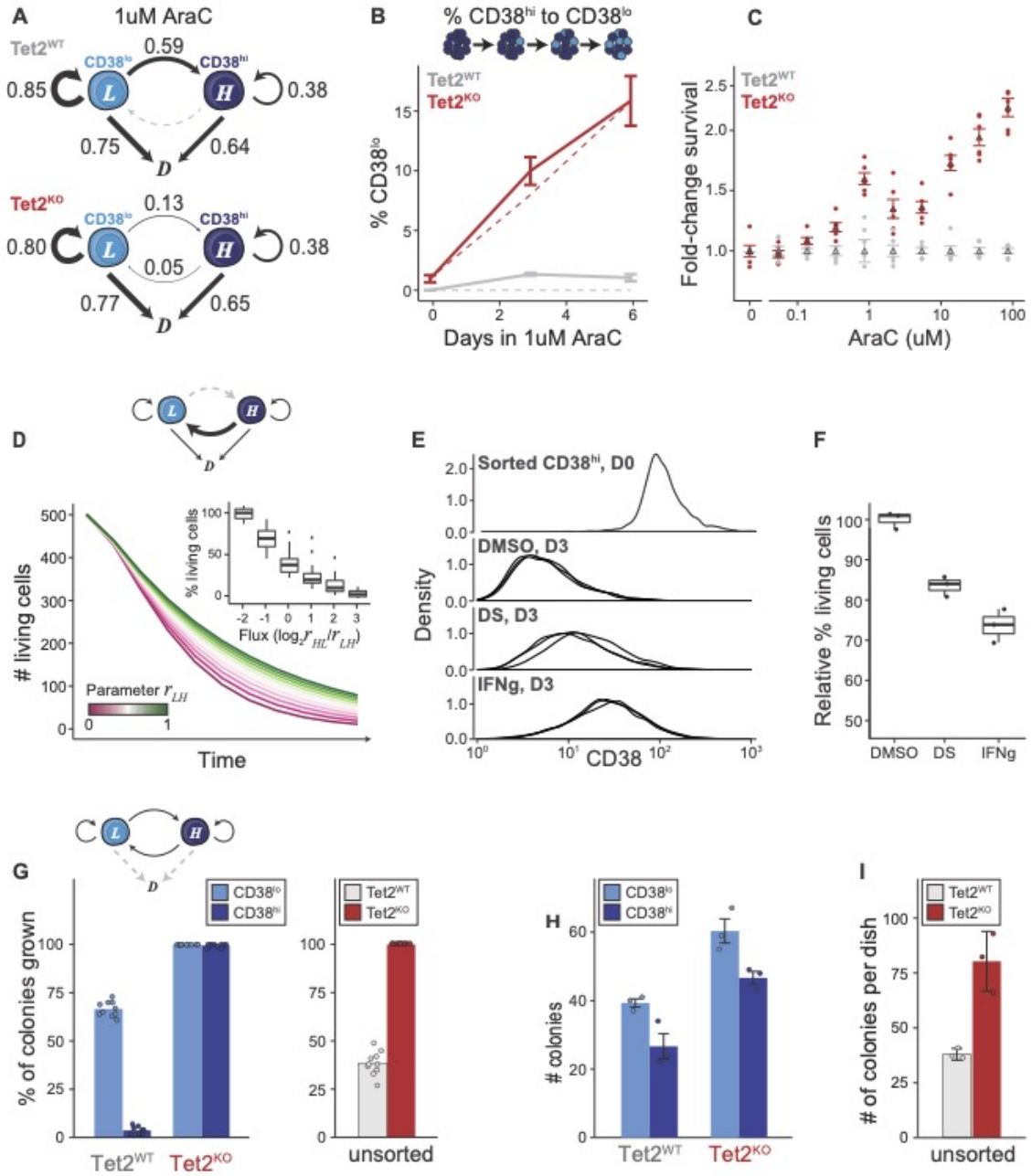

Figure 2.4: Altered cell-state dynamics enable longer-term drug survival in chemotherapy
(A) Schematic illustration of model of cell-state transition, cell proliferation, and cell death in 1μM AraC treatment. Numbers are representative estimated rates of transition, proliferation, and death (see Figure A.10). (B) The percent of $CD38^{lo}$ cells in an originally pure population of sorted $CD38^{hi}$ cells in 1μM AraC after 0, 3, and 6 days (mean ± s.e., n=3). Dashed lines represent predictions from the model. (C) The viability of cells treated with 72h of varying concentrations of AraC were assessed by CellTiter-Glo. Shown is fold-change viability relative to $TET2^{WT}$ cells as a function of chemotherapy concentration (mean ± s.e., n=5). (D) Model predictions for cell drug survival given increasing values of the transition parameter r_{HL} for a toy model that fulfills Inequality 1. Plotting survival as a function of the ratio of r_{HL}/r_{LH} (inset). (E) $CD38^{hi}$ cells were isolated by FACS, then incubated with or without effector for 3 days in the presence of

1μM AraC. Shown are distributions of sorted $TET2^{KO}$ $CD38^{hi}$ cells after 3 days in DMSO, DS, or IFNg treatment (n=3, $TET2^{WT}$ in Figure A.12). (F) Sorted $CD38^{hi}$ cells were subjected to 1μM AraC treatment with or without effector. The percent of living cells in the population relative to control after 3 days is shown on the y-axis (mean ± s.e., n=3). (G) Model predictions for colony outgrowth from a small population given sorted $CD38^{lo}$ and $CD38^{hi}$ (left) or unsorted cells (right; see Methods, mean of 10 runs). (H) Number of colonies in a methylcellulose colony forming unit assay for $TET2^{WT}$ and $TET2^{KO}$ cells after seeding equal numbers and 14d incubation (mean ± s.e., n=3). (I) Number of colonies in methylcellulose assay $TET2^{WT}$ and $TET2^{KO}$ cells sorted for CD38 expression (mean ± s.e., n=3).

Methods

Cell-state transition model

We modeled a cell system with m states as a linear, homogenous system of ODEs. In matrix form this is written as

$$\frac{dv}{dt} = Mv$$

where v is cell-state vector of length m, and M is a $m\ x\ m$ matrix representing cell-state transitions. We do not assume the matrix M is stochastic, which allows expansion or contraction of the total number of cells in the system—reflecting cell division or drug sensitivity—to occur. In our study, we considered two states: L (CD38^{lo}) or H (CD38^{hi}):

$$\frac{d}{dt}\begin{bmatrix} L \\ H \end{bmatrix} = \begin{bmatrix} (g_L - r_{LH} - d_L) & r_{HL} \\ r_{LH} & (g_H - r_{HL} - d_H) \end{bmatrix}\begin{bmatrix} L \\ H \end{bmatrix}$$

Where g_L and g_H are the growth rates and d_L and d_H the death rates of cell-states L and H, respectively. Here, r_{LH} is the transition rate from L to H, and r_{HL} is the transition rate from H to L. All matrix parameters are considered to be real, non-negative numbers. If we define

$$\alpha = -g_L + r_{LH} + r_{HL} - g_H + d_L + d_H$$

$$\beta = \sqrt{\alpha^2 + 4r_{LH}(g_H - d_H) + 4r_{HL}(g_L - d_L) + 4(g_H d_L + g_L d_H) - 4(g_L g_H + d_L d_H)}$$

then the eigenvalues of M are

$$\lambda_1 = \frac{1}{2}(-\alpha - \beta)$$

$$\lambda_2 = \frac{1}{2}(-\alpha + \beta)$$

with eigenvectors

$$\overrightarrow{v_1} = \begin{bmatrix} -\dfrac{-g_L + r_{LH} - r_{HL} + g_H + d_L - d_H + \beta}{2r_{HL}} \\ 1 \end{bmatrix}$$

$$\overrightarrow{v_2} = \begin{bmatrix} -\dfrac{-g_L + r_{LH} - r_{HL} + g_H + d_L - d_H - \beta}{2r_{HL}} \\ 1 \end{bmatrix}$$

For convenience, we define the difference in production between the two states

$$P \equiv (g_L - d_L) - (g_H - d_H)$$

To find the regime where changing r_{HL} (asymptotically) increases survival rate, we focus on the larger of the other two eigenvalues, namely λ_2. Taking the derivative of λ_2 with respect to the state-switching parameters gives:

$$\frac{d\lambda_2}{dr_{HL}} = \frac{1}{2}\left(\frac{r_{LH} + r_{HL} + P}{\beta} - 1\right)$$

$$\frac{d\lambda_2}{dr_{LH}} = \frac{1}{2}\left(\frac{r_{LH} + r_{HL} - P}{\beta} - 1\right)$$

Using Wolfram|Alpha, we find: a) $\frac{d\lambda_2}{dr_{HL}} > 0$ holds if $P > 0$ and $r_{HL} > -\left(\sqrt{r_{LH}} - \sqrt{P}\right)^2$; and b) $\frac{d\lambda_2}{dr_{LH}} \leq 0$ holds if either (i) $P > 0$, $0 \leq r_{LH} < P$, and $r_{HL} > -\left(\sqrt{r_{LH}} - \sqrt{P}\right)^2$, or if (ii) $P > 0$ and $r_{LH} \geq P$. In our model, we only consider non-negative, real-valued parameter values. Thus, as long as the inequality $P > 0$ holds, $\frac{d\lambda_2}{dr_{HL}} > 0$ and $\frac{d\lambda_2}{dr_{LH}} \leq 0$.

A two-state system with constant switching rates can equilibrate to a constant fraction of states (a steady state) provided the cell population does not rapidly crash. This is possible because flux into and out of each state can be balanced and is why observations such as Figure 2.2 (where $TET2^{KO}$ cells appear to not switch and instead retain a stem-like state) can still be an accurate reflection of cells in constant flux. However, cell proliferation, death, and switching rates cannot be disentangled directly from bulk measurements. Therefore, we estimated these rates by fitting our model to time course data of sorted $TET2^{KO}$ and WT populations. To find best fit parameters for this system, cell-state compositions from the flow cytometry experiments shown in Figure A.8 or Figure A.9 (using data from both $CD38^{hi}$- and $CD38^{lo}$-sorted subpopulations, described below) were used as input into the model. Prediction error (the sum of the differences between the predicted and measured number of cells in each cell-state at each timepoint) was minimized with fminsearch in MATLAB (version R2019a) with MaxIter of 5000. To eliminate "solutions" that did not converge or had poor fit, prediction error for 1000 sets of random parameter values was

calculated to form a null distribution for each condition, and the 0.005-quantile for this null distribution was used as an upper limit for reasonable error values (Figure A.10). To establish the probable range of transition parameters, best-fit parameters for 1000 random initializations (x0 in the fminsearch function) were found and plotted (Figure A.10). Representative values for fit parameters for KG1 TET2^{WT} and TET2^{KO} cells are shown in Figure 2.4. Median values for fit parameters or fit parameter ratios from converged solutions are shown in Figure A.11 and Figure A.13. Code is available online at https://github.com/AltschulerWu-Lab/TET2-dynamics.

Solving the matrix ODE gave

$$\begin{bmatrix} L \\ H \end{bmatrix} = \begin{bmatrix} \vec{v}_1^1 & \vec{v}_1^2 \\ \vec{v}_2^1 & \vec{v}_2^2 \end{bmatrix} \begin{bmatrix} Ae^{\lambda_1 t} \\ Be^{\lambda_2 t} \end{bmatrix}$$

After solving for the constants A and B, model predictions were made using representative best fit parameters and initial conditions observed in the experiment (e.g. number of CD38^{hi} and CD38^{lo} cells in day 0 sorted cell populations; method for calling CD38$^{lo/hi}$ described below).

For colony forming simulations, 100 colonies were "seeded" with 10 individual cells as colony founders with 50% in the L state and 50% in H unless otherwise noted. To simulate growth without drug treatment, rates were estimated by fitting the model to flow cytometry experimental data (described below) without AraC treatment. The Gillespie algorithm was implemented to find the number of cells in the L or H states at time t. Cells "grew" for 10 time-steps before the number of cells in each colony was counted. Colonies with more than 100 cells were considered "grown out".

Cell lines, reagents, and cell culture

KG1 cells and Thp1/Thp1 TET2^{KD} cells were generous gifts from the Shannon lab at UCSF and Levine lab at MSKCC, respectively. Cells were cultured in ATCC-recommended media and incubated at 37°C with 5% CO_2. All cell lines were maintained in 75ml or 250ml Suspension Culture flasks (CellTreat, Pepperell, MA) between 0.5e6 and 1e6 cells/ml.

To make TET2^{KO} cells in the KG1 cell line, guides targeting exon 3 of TET2 (an exon with frequent indel mutations in patients with MDS; Smith *et al*., 2010) were designed in Benchling (5'-

CACCGAGGCCAATTAAGGTGGAACC-3'). pSpCas9(BB)-2A-Puro (PX459) V2.0 was a gift from Feng Zhang (Addgene plasmid #62988). Guides were cloned into PX459 using BbsI sites, and vectors transfected into KG1 cells with the Cell Line Nucleofector Kit R (Lonza Group, Basel, Switzerland) with protocol V-001 and according to the manufacturer's protocol. After 48h, FITC-positive cells were isolated via FACS and expanded for 10 days. The TET2 mutation created a frameshift mutation causing truncation (Figure A.1) upstream of the conserved catalytic domain of TET2, consistent with tumors observed in patients (Weissman *et al*., 2012).

Prior to addition to cell culture media, Cytosine Beta-D-Arabinofuranoside (Sigma-Aldrich, St. Louis, MO) and 2-Deoxy-D-glucose (Sigma-Aldrich) were dissolved in H_2O, CCCP (Sigma-Aldrich), Disulfiram (Sigma-Aldrich), Etacrynic acid (Sigma-Aldrich), and Torin1 (Cell Signaling Technology, Danvers, Massachusetts) were dissolved in DMSO, and IFNg (Peprotech, Rocky Hill, NJ) was dissolved in culture media to 1000x the desired final concentration. Final DMSO (or other diluent) concentration was always 0.1%.

To generate conditioned media, 5e6 KG1 TET2 WT or KO cells were seeded in fresh growth media and grown for 24h. Samples were centrifuged at 300g for 3 minutes, and fresh growth media was added 1:1 to the supernatant prior to filtering with 0.4um Steriflip units (Millipore, Burlington, MA). Cells were treated for 6 days with conditioned media (unless cut short due to Covid-19 shelter-in-place). Conditioned media was made fresh for each timepoint.

DNA methylation profiling

DNA methylation was profiled using Illumina's Infinium MethylationEPIC BeadChip (Illumina, San Diego, CA). DNA of technical replicates for each condition was extracted using the Zymo Quick-DNA kit (Zymo Research, Irvine, CA, KG1 n=6 per condition, Thp1 n=2 per condition). Bisulfite conversion, nanodrop quantitation, array scanning, and normalization was performed by the Vincent J. Coates Genomics Sequencing Laboratory at UC Berkeley. Differential methylation analysis was performed with ChAMP (package version 2.14.0, Tian *et al*., 2017) in R (version 3.6.0; R Core Team, 2019). Differential

methylated region analysis was performed with DMRcate (version 1.20, Peters *et al*., 2015), and visualization done with Gviz (version 1.28, Hahne *et al*., 2016).

mRNA-seq

RNA extraction of technical replicates was performed using the Lexogen SPLIT RNA extraction kit (Lexogen, Vienna, Austria, n=6 per condition), and libraries were prepared using the QuantSeq 3' mRNA-Seq Library Prep Kit FWD for Illumina. Samples were quantified with Invitrogen Qubit (Invitrogen, Carlsbad, CA) prior to pooling, and library size and integrity was confirmed using the Agilent Bioanalyzer with the high-sensitivity DNA kit (Agilent, Santa Clara, CA). RNA sequencing was performed using 50bp single-end sequencing on the Illumina HiSeq 4000 in the Center for Advanced Technology at UC San Francisco. A PhiX control library was used as an in-run control, spiked in at 5%. Reads were mapped and counted using the Integrated QuantSeq data analysis pipeline on Bluebee Platform (Bluebee, Rijswijk, Netherlands). Briefly, reads were trimmed with BBDuk, aligned to human GRCh38 with STAR, and counted with HTSeq-count.

Gene filtering and differential expression analysis was performed in R. Genes were filtered by count such that all genes had 3 or more samples with 10 or more counts. Differential expression analysis was then performed using DESeq2 (version 1.24, Love *et al*., 2014) on gene counts. Genes that were found to be significantly differentially expressed with an absolute Log_2fold-change in expression > 1.5X in a paired analysis of untreated cell lines were submitted to Enrichr (Kuleshov *et al*., 2016) for enrichment analysis.

TET2 expression analysis

Expression profiles for the following datasets were queried from the Cancer Genomics Data Server (CGDS) using cdgsr (version 1.3.0, Cerami *et al*., 2012). Our results are, in part, based upon data generated by The Cancer Genome Atlas managed by the NCI and NHGRI. Information about TCGA can be found at http://cancergenome.nih.gov. For the datasets TCGA-LAML (dbGaP accession phs000178.v1.p1] and AML-OHSU (Tyner *et al*., 2018), genes with coefficient of variation < 0.1 across samples were dropped from subsequent analysis. Pearson correlation calculations were performed in R (version 3.6.0). Genes with

expression that strongly correlated (cor > 0.45) or anti-correlated (cor < -0.45) with TET2 expression were submitted to Enrichr (Kuleshov *et al*., 2016) for enrichment analysis.

Methylation age and stemness profiles

Quantile normalized values from the aforementioned ChAMP analysis were used as input to the DNA methylation age calculator in R as described in the tutorial (horvath.genetics.ucla.edu) (Horvath, 2013). For visualization purposes, results for $TET2^{WT}$ and $TET2^{KO}$ cells were normalized to $TET2^{WT}$.

For the "stem-like" epigenetic signatures, we used a publicly available dataset of DNA methylation profiles from normal hematopoietic progenitors and leukemic cells sorted by CD34/CD38 expression (Jung *et al*., 2015). Linear discriminant analysis (LDA) was applied using the MASS package (version 7.3-51.4, Venables *et al*., 2002) in R to identify an optimal transform that increased separation of DNA methylation profiles across three key reference populations: normal HSCs, $CD34^{+}CD38^{-}$ putative LSCs, and $CD34^{-}$ leukemia cells. The rest of the public dataset, as well as methylation data from the paired $TET2^{WT}$/$TET2^{KO}$ cell lines, was projected into the lower-dimensional LDA space.

Immunofluorescence staining for flow cytometry

Cells were pelleted and washed with wash buffer (HBSS + 1% BSA, filtered with a 50ml Steriflip unit (Millipore)) prior to 30m incubation in Fc Receptor Blocker (Innovex Biosciences, Richmond, CA) on ice in the dark. Cells were washed twice before resuspension in wash buffer containing conjugated antibodies for flow cytometry at their recommended concentrations (CD34-PE 555822, CD38-PE-Cy5 555461, Becton Dickinson, Franklin Lakes, NJ), and incubation for 30m on ice in the dark. Cells were washed three times and resuspended in 350µL wash buffer before measurement with flow cytometry at a flow rate of 9.0. Doublets were called based on gates drawn for FSC-A and FSC-W, and dead cells were counted based on gates drawn in FSC-A and SSC-A. All flow cytometry or FACS was performed on the Aria IIu in the Center for Advanced Technologies at UC San Francisco.

External marker tracking by flow cytometry

For Figure 2.2, Figure A.8 and Figure A.9, equal numbers of $CD38^{hi}$ and $CD38^{lo}$ subpopulations of $TET2^{WT}$ and $TET2^{KO}$ cells were isolated by FACS (gate in Figure 2.2, set such that <0.5% of unstained

cells would be counted as $CD38^{hi}$; see Figure A.5) and seeded separately in a round-bottom 96-well plate (CELLSTAR) at 1e5 cells/ml. For Figure A.8, equal numbers of cells from the "tails" (bottom/top 5%) of the CD38 distributions of stained, unsorted $TET2^{WT}$ and $TET2^{KO}$ cells (see Figure A.8) were isolated by FACS and seeded separately in a round-bottom 96-well plate (CELLSTAR) at 1e5 cells/ml. For day 0 timepoints, aliquots of freshly sorted cells were reflowed and recorded. For Figure 2.2 and Figure A.8, cells were treated with 0, 1, or 4μM AraC in technical triplicates. No drug was applied to the technical replicates shown in Figure A.8 and Figure A.9. Plates were covered with Breathe-Easy sealing membranes (Sigma-Aldrich) before incubation. For longer timepoints, media (with or without drug, as relevant) was exchanged every 3 days. At each timepoint, all cells in each replicate well were washed and stained for CD34 and CD38 (described above) and re-profiled by flow cytometry. To account for day-to-day variations in Aria IIu laser power or other settings, raw fluorescence intensity for each channel was normalized by the median intensity of unsorted, unstained controls of the appropriate cell line (Figure A.5) for each timepoint prior to analysis. For all experiments, $CD38^{hi}$ and $CD38^{lo}$ cell counts for analysis and model-fitting were called based on the same sorting threshold shown in Figure 2.2 (set such that <0.5% of unstained cells would be counted as $CD38^{hi}$; see Figure A.5). The number of living cells in each cell state served as input to the model to fit parameters for the cell-state transition matrix in both cell lines for both treated and untreated conditions (described above).

MethoCult methylcellulose colony forming assay

Cells were counted with the TC20 automatic cell counter (Bio-Rad) with Trypan blue prior to plating in triplicate as technical replicates in MethoCult H4034 Optimum (Stemcell Technologies, Vancouver, CAN) in 35mm cell culture dishes (Eppendorf, Hamburg, DE) according to the manufacturer's instructions. Cells were incubated for 2 weeks at 37°C with 5% CO_2. To maintain humidity, 4 MethoCult dishes were incubated amongst three lidless 35mm dishes each filled with 3ml sterile ddH_2O inside a 150mm glass petri dish with a lid (Corning, Corning, NY). Water was replenished every 3 days. Colonies were enumerated according to the manufacturer's instructions. Results shown are representative results from two independent experiments.

Cell viability assays

For sorted population growth or effector viability assays (Figure A.7 and Figure A.10), $TET2^{WT}$ and $TET2^{KO}$ cells were stained for CD34 and CD38 (described above) and segregated by CD38 expression with FACS as shown (Figure A.7 and Figure 2.2, respectively). For day 0 timepoints, aliquots of freshly sorted cells were reflowed and recorded. Sorted cells were plated in round-bottom 96-well plates (CELLSTAR, Dallas, TX) at 1e5 cells/ml with replicates for each timepoint. For treated conditions, AraC and effectors were diluted and added as described above. Plates were covered with Breathe-Easy sealing membranes (Sigma-Aldrich) before incubation at 37°C with 5% CO_2. For longer timepoints, media (with or without drug, as relevant) was exchanged every 3 days. At each timepoint, all cells in each replicate well were washed, stained, and measured with the Aria IIu for the first 2000 events. The density of living cells (# of cells per ml, Figure A.7) was calculated as the number of living cells observed per sample divided by the total length of time the sample took to reach 2000 events, and divided by the flow rate (~90 uL/minute for flow rate 9.0 on the Aria IIu).

For drug viability assays (Figure 2.4, Figure A.8), cells were plated as technical replicates at 1e5 cells/ml and treated with drug for 72h. For readout, plates were allowed to cool to room temperature before combining well-mixed cells and cell media 1:1 with room temperature CellTiter-Glo 2.0 (Promega, Madison, WI) in opaque white tissue culture plates (Corning). Reactions were allowed to proceed according to the manufacturer's protocol, and luminescence was read out with the Biotek H4 plate reader (BioTek, Winooski, VT) in the Center for Advanced Technology at UC San Francisco. Results shown are representative results from three independent experiments.

RT-qPCR and PCR

For measuring TET2 expression, RNA was extracted from three technical replicates with the Direct-zol RNA Miniprep kit (Zymo Research) according to manufacturer's instructions with TRI Reagent (Thermo Fisher Scientific, Waltham, MA). For measuring CD38, ITGAM, CORO1A, TAL1, and HOXA5 expression, cells were stained for CD34 and CD38 (described above) and fractionated by CD38 expression via FACS (gates shown in Figure A.6). RNA was extracted from sorted cells (2 replicates of 300k cells)

with the RNeasy Plus Mini Kit (QIAGEN, Hilden, DE) according to manufacturer's instructions. Reverse transcription with iScript Reverse Transcription Supermix (Bio-Rad, Hercules, CA) was followed by qPCR with DyNAmo Flash SYBR Green qPCR kit (Thermo Fisher Scientific) according to the recommended protocol on the BioRad CFX Connect in the Center for Advanced Technologies at UC San Francisco. The set of genes measured was composed of genes of interest from the bulk RNAseq experiment (Figure A.2) and relevant gene sets in the literature (GSEA: gal_leukemic_stem_cell (Gal *et al*., 2006), gentles_leukemic_stem_cell (Gentles *et al*., 2010), eppert_ce_hsc_lsc (Eppert *et al*., 2011)). Primers were obtained from IDT (San Jose, CA). Primer sequences for human RUNX1, TET2, HOXA5, and GAPDH were as previously described (Fujita *et al*., 2001; Cimmino *et al*., 2017; McLaughlin-Drubin *et al*., 2011). Primer sequences for CD38, CORO1A, ITGAM, and TAL1 were generated with the Primer Design Tool in Benchling.

Data availability

RNA sequencing, DNA methylation, flow cytometry, and other data that support the findings of this study is publicly available in Mendeley Data along with code to generate all figures (doi: 10.17632/xmvz47rpg6.1). Code to generate parameter fits or run the Gillespie algorithm is available on GitHub at https://github.com/AltschulerWu-Lab/TET2-dynamics.

References

Amici SA, Young NA, Narvaez-Miranda J, Jablonski KA, Arcos J, Rosas L, Papenfuss TL, Torrelles JB, Jarjour WN, and Guerau-de-Arellano M. (2018). CD38 Is Robustly Induced in Human Macrophages and Monocytes in Inflammatory Conditions. Front in Immunology, 9, 1593.

Asmar F, Punj V, Christensen J, Pedersen MT, Nielsen AB, Hother C, Ralfkiaer U, Brown P, Ralfkiaer E, *et al*. (2013). Genome-wide profiling identifies a DNA methylation signature that associates with TET2 mutations in diffuse large B-cell lymphoma. Haematologica, 98(12), 1912–1920.

Bai X, Fisher DE, Flaherty, KT. (2019). Cell-state dynamics and therapeutic resistance in melanoma from the perspective of MITF and IFNg pathways. Nat Reviews Clinical Oncology, 16, 549-562.

Brauner A, Fridman O, Gefen O, Balaban, NQ. (2016). DistinguA-ishing between resistance, tolerance and persistence to antibiotic treatment. Nature Reviews Microbiology, 14(5), 320–330.

Busque L, Patel J, Figueroa M, *et al*. (2012). Recurrent somatic TET2 mutations in normal elderly individuals with clonal hematopoiesis. Nature Genetics 44, 1179–1181.

Cancer Genome Atlas Research Network. (2013). Genomic and epigenomic landscapes of adult de novo acute myeloid leukemia. The New England Journal of Medicine, 368(22), 2059–2074.

Cerami E, Gao J, Dogrusoz U, Gross BE, Sumer SO, Aksoy BA, Jacobsen A, Byrne CJ, Heuer ML, Larsson E, *et al*. (2012). The cBio Cancer Genomics Portal: An Open Platform for Exploring Multidimensional Cancer Genomics Data. Cancer Discovery, 2(5), 401-404.

Chan KT, Creed SJ & Bear JE. (2012). Coronin family of actin regulators. 21(8), 481–488.

Cheson BD, Bennett JM, Kopecky KJ, Büchner T, Willman CL, Estey EH, Schiffer CA, Doehner H, Tallman MS, Lister TA, *et al*. (2003). Revised recommendations of the International Working Group for Diagnosis, Standardization of Response Criteria, Treatment Outcomes, and Reporting Standards for Therapeutic Trials in Acute Myeloid Leukemia. Journal of Clinical Oncology, 21(24), 4642–4649.

Cimmino L, Dolgalev I, Wang Y, Yoshimi A, Martin GH, Wang J, Ng V, Xia B, Witkowski MT, Mitchell-flack M, *et al*. (2017). Restoration of TET2 Function Blocks Aberrant Self-Renewal and Leukemia Progression. Cell, 170(6), 1079–1095.

Corces-Zimmerman MR, Hong WJ, Weissman IL, Medeiros BC, Majeti R. (2014). Preleukemic mutations in human acute myeloid leukemia affect epigenetic regulators and persist in remission. Proceedings of the National Academy of Sciences, 111(7), 2548–2553.

Costello RT, Mallet F, Gaugler B, Sainty D, Arnoulet C, Gastaut J, Olive D. (2000). Human acute myeloid leukemia CD34+CD38- progenitor cells have decreased sensitivity to chemotherapy and Fas-induced apoptosis, reduced immunogenicity, and impaired dendritic cell transformation capacities. Cancer Research, 60(16), 4403-4411.

Ding L, Ley TJ, Larson DE, Miller CA, Koboldt DC, Welch JS, Ritchey JK, Young MA, Lamprecht T, McLellan MD, *et al*. (2012). Clonal evolution in relapsed acute myeloid leukaemia revealed by whole-genome sequencing. Nature, 481(7382), 506–510.

Eppert K, Takenaka K, Lechman ER, Waldron L, Nilsson B, van Galen P, Metzeler KH, Poeppl A, Ling V, Beyene J, *et al*. (2011). Stem cell gene expression programs influence clinical outcome in human leukemia. Nature Medicine, 17(9), 1086-93.

Feinberg AP, Koldobskiy MA, Göndör A. (2016). Epigenetic modulators, modifiers and mediators in cancer aetiology and progression. Nature Reviews Genetics, 17, 284-299.

Flavahan WA, Gaskell E, Bernstein, BE. (2017). Epigenetic plasticity and the hallmarks of cancer. Science, 357(6348).

Fujita Y, Nishimura M, Taniwaki M, Abe T, Okuda T. (2001). Identification of an alternatively spliced form of the mouse AML1/RUNX1 gene transcript AML1c and its expression in early hematopoietic development. Biochem and Biophys Research Comms, 281(5), 1248–1255.

Gal H, Amariglio N, Trakhtenbrot L, Jacob-Hirsh J, Margalit O, Avigdor A, Nagler A, Tavor S, Ein-Dor L, Lapidot T, *et al*. (2006). Gene expression profiles of AML derived stem cells; similarity to hematopoietic stem cells. Leukemia, 20(12), 2147-54.

Gentles AJ, Plevritis SK, Majeti R, Alizadeh, AA. (2010). Association of a leukemic stem cell gene expression signature with clinical outcomes in AML. JAMA, 304(24), 2706-2715.

Gerber JM, Smith BD, Ngwang B, Zhang H, Vala MS, Morsberger L, Galkin S, Collector MI, Perkins B, Levis MJ, *et al*. (2012). A clinically relevant population of leukemic CD34+CD38- cells in acute myeloid leukemia. Blood, 119(15), 3571-3577.

Girgis HS, Harris K, and Tavazoie S. (2012). Large mutational target size for rapid emergence of bacterial persistence. Proceedings of the National Academy of Sciences of the United States of America, 109(31), 12740–12745.

Gupta PB, Fillmore CM, Jiang G, Shapira SD, Tao K, Kuperwasser C, Lander ES. (2011). Stochastic state transitions give rise to phenotypic equilibrium in populations of cancer cells. Cell, 146(4), 633–644.

Hahne F & Ivanek R. (2016). Statistical Genomics: Methods and Protocols. In Visualizing Genomic Data Using Gviz and Bioconductor (335-351). New York, NY: Springer New York.

Hinohara K, Wu HJ, Vigneau S, McDonald TO, Igarashi KJ, Yamamoto KN, Madsen T, Fassl A, Egri SB, Papanastasiou M, *et al*. (2018). KDM5 histone demethylase activity links cellular transcriptomic heterogeneity to therapeutic resistance. Cancer Cell, 34(6), 939-953.e9.

Hirsch CM, Nazha A, Kneen K, Abazeed ME, Meggendorfer M, Przychodzen BP, Nadarajah N, Adema V, Nagata Y, Goyal A, *et al*. (2018). Consequences of mutant TET2 on clonality and subclonal hierarchy. Leukemia, 32(8), 1751–1761.

Horvath S. (2013). DNA methylation age of human tissues and cell types. Genome Biology, R115.

Jordan NV, Bardia A, Wittner BS, Benes C, Ligorio M, Zheng Y, Yu M, Sundaresan TK, Licausi JA, Desai R, *et al*. (2016). HER2 expression identifies dynamic functional states within circulating breast cancer cells. Nature, 537(7618), 102–106.

Jung N, Dai B, Gentles AJ, Majeti R, and Feinberg AP. (2015). An LSC epigenetic signature is largely mutation independent and implicates the HOXA cluster in AML pathogenesis. Nature Communications, 6, 8489.

Kanehisa M. (2000). KEGG: Kyoto Encyclopedia of Genes and Genomes. Nucleic Acids Research, 28(1), 27–30.

Kanehisa M, Sato Y, Furumichi M, Morishima K & Tanabe M. (2019). New approach for understanding genome variations in KEGG. Nucleic Acids Research, 47(D1), D590–D595.

Kuleshov MV, Jones MR, Rouillard AD, Fernandez NF, Duan Q, Wang Z, Koplev S, Jenkins SL, Jagodnik KM, Lachmann A, *et al*. (2016). Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. Nucleic Acids Research, 44(W1), W90–7.

Kunimoto H, Meydan C, Nazir A, Whitfield J, Shank K, Rapaport F, Maher R, Pronier E, Meyer SC, *et al*. (2018). Cooperative Epigenetic Remodeling by TET2 Loss and NRAS Mutation Drives Myeloid Transformation and MEK Inhibitor Sensitivity. Cancer Cell, 33(1), 44–59.e8.

La Manno G, Soldatov R, Zeisel A, Braun E, Hochgerner H, Petukhov V, Lidschreiber K, Kastriti ME, Lönnerberg P, Furlan A, *et al*. (2018). RNA velocity of single cells. Nature, 560(7719), 494–498.

Love MI, Huber W, Anders S. (2014). Moderated Estimation of Fold Change and Dispersion for RNA-seq Data with DESeq2. Genome Biology, 15, 550.

Ma C, Wei S, and Song Y. (2011). T790M and acquired resistance of EGFR TKI: a literature review of clinical reports. Journal of Thoracic Disease, 3(1), 10–18.

McLaughlin-Drubin ME, Crum CP, and Münger K. (2011). Human papillomavirus E7 oncoprotein induces KDM6A and KDM6B histone demethylase expression and causes epigenetic reprogramming. Proceedings of the National Academy of Sciences of the United States of America, 108(5), 2130-2135.

Metzeler KH, Maharry K, Radmacher MD, Mrózek K, Margeson D, Becker H, Curfman J, Holland KB, Schwind S, Whitman SP, *et al*. (2011). TET2 Mutations Improve the New European LeukemiaNet Risk Classification of Acute Myeloid Leukemia: A Cancer and Leukemia Group B Study. Journal of Clinical Oncology, 1373–1381.

Moran-Crusio K, Reavie L, Shih A, Abdel-Wahab O, Ndiaye-Lobry D, Lobry C, Figueroa ME, Vasanthakumar A, Patel J, Zhao X, *et al*. (2011). TET2 loss leads to increased hematopoietic stem cell self-renewal and myeloid transformation. Cancer Cell, 20(1), 11–24.

Nibourel O, Kosmider O, Cheok M, Boissel N, Renneville A, Philippe N, Dombret H, Dreyfus F, Quesnel B, Geffroy S, *et al*. (2010). Incidence and prognostic value of TET2 alterations in de novo acute myeloid leukemia achieving complete remission. Blood, 116(7), 1132–1135.

Nishioka C, Ikezoe T, Furihata M, Yang J, Serada S, Naka T, Nobumoto A, Kataoka S, Tsuda M, Udaka K, and Yokoyama A. (2013). CD34+/CD38− acute myelogenous leukemia cells aberrantly express CD82 which regulates adhesion and survival of leukemia stem cells. International journal of cancer. Journal International du Cancer, 132(9), 2006–2019.

Peters TJ, Buckley MJ, Statham AL, Pidsley R, Samaras K, Lord RV, Clark SJ, Molloy, PL. (2015). De novo identification of differentially methylated regions in the human genome. Epigenetics & Chromatin, 8, 6.

Potter N, Miraki-Moud F, Ermini L, *et al*. (2019). Single cell analysis of clonal architecture in acute myeloid leukaemia. Leukemia 33, 1113–1123.

R Core Team (2019). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Ran D, Shia WJ, Lo MC, Fan JB, Knorr DA, Ferrell PI, Ye Z, Yan M, Cheng L, Kaufman DS, *et al*. (2013). RUNX1a enhances hematopoietic lineage commitment from human embryonic stem cells and inducible pluripotent stem cells. Blood, 121(15), 2882–2890.

Rasmussen KD, Jia G, Johansen JV, Pedersen MT, Rapin N, Bagger FO, Porse BT, Bernard OA, Christensen J, and Helin K. (2015). Loss of TET2 in hematopoietic cells leads to DNA hypermethylation of active enhancers and induction of leukemogenesis. Genes & Development, 29(9), 910–922.

Rothenberg-Thurley M, Amler S, Goerlich D, Köhnke T, Konstandin NP, Schneider S, Sauerland MC, Herold T, Hubmann M, Ksienzyk B, *et al*. (2018). Persistence of pre-leukemic clones during first remission and risk of relapse in acute myeloid leukemia. Leukemia, 32(7), 1598–1608.

Sanchez-Vega F, Mina M, Armenia J, Chatila WK, Luna A, La KC, Dmitriadoy S, Liu DL, Kantheti HS, Saghafinia S, *et al*. (2018). Oncogenic signaling pathways in The Cancer Genome Atlas. Cell, 173(2), 321-337.e10.

Shih AH, Jiang Y, Meydan C, Shank K, Pandey S, Barreyro L, Antony-Debre I, Viale A, Socci N, Sun Y, *et al*. (2015). Mutational cooperativity linked to combinatorial epigenetic gain of function in acute myeloid leukemia. Cancer Cell, 27(4), 502–515.

Smith AE, Mohamedali AM, Kulasekararaj A, Lim Z, Gäken J, Lea NC, Przychodzen B, Mian SA, Nasser EE, Shooter C, *et al*. (2010). Next-generation sequencing of the TET2 gene in 355 MDS and CMML patients reveals low-abundance mutant clones with early origins, but indicates no definite prognostic value. Blood, 116(19), 3923–3932.

Su Y, Wei W, Robert L, Xue M, Tsoi J, Garcia-Diaz A, Homet Moreno B, Kim J, Ng RH, Lee JW, *et al*. (2017). Single-cell analysis resolves the cell state transition and signaling dynamics associated with melanoma drug-induced resistance. Proceedings of the National Academy of Sciences of the United States of America, 114(52), 13679–13684.

Tian Y, Morris TJ, Webster AP, Yang Z, Beck S, Andrew F, Teschendorff AE. (2017). ChAMP: updated methylation analysis pipeline for Illumina BeadChips. Bioinformatics, 33(24), 3982-3984.

Tyner JW, Tognon CE, Bottomly D, Wilmot B, Kurtz SE, Savage SL, Long N, Schultz AR, *et al*. (2018). Functional genomic landscape of acute myeloid leukaemia. Nature, 562(7728), 526–531.

Vagapova ER, Spirin PV, Lebedev TD & Prassolov VS. (2018). The role of TAL1 in hematopoiesis and leukemogenesis. Acta Naturae, 10(1), 15–23.

Venables, WN & Ripley, BD. (2002). Modern Applied Statistics with S.

Weissmann S, Alpermann T, Grossmann V, Kowarsch A, Nadarajah N, Eder C, Dicker F, Fasan A, Haferlach C, Haferlach T, Kern W, Schnittger S & Kohlmann A. (2012). Landscape of TET2 mutations in acute myeloid leukemia. Leukemia, 26(5), 934–942.

Wolfram Research, Inc., Wolfram|Alpha. www.wolframalpha.com. (Accessed: 22 September 2019).

Xu B, Wang S, Li R, Chen K, He L, Deng M, Kannappan V, Zha J, Dong H, and Wang W. (2017). Disulfiram/copper selectively eradicates AML leukemia stem cells *in vitro* and *in vivo* by simultaneous induction of ROS-JNK and inhibition of NF-κB and Nrf2. Cell Death & Disease, 8(5), e2797.

Yamazaki J, Jelinek J, Lu Y, Cesaroni M, Madzo J, Neumann F, He R, Taby R, Vasanthakumar A, Macrae T, *et al*. (2015). TET2 Mutations Affect Non-CpG Island DNA Methylation at Enhancers and Transcription Factor-Binding Sites in Chronic Myelomonocytic Leukemia. Cancer Research, 75(14), 2833–2843.

Zeijlemaker W, Grob T, Meijer R, Hanekamp D, Kelder A, Carbaat-Ham JC, Oussoren-Brockhoff, YJM, Snel AN, Velhuizen D, *et al*. (2018). $CD34^{+}CD38^{-}$ leukemic stem cell frequency to predict outcome in acute myeloid leukemia. Leukemia, 33, 1102-1112.

Chapter 3

Metabolic effects of TET2 loss

Main contributing authors

Leanna Morinishi[1], Karl Kochanowski[2], Maike Roth[2], Maren Diether[3], Janice Goh[4], Mollie E. Hansen[5], Lani F. Wu[2] & Steven J. Altschuler[2]

Affiliations

[1] Bioinformatics Graduate Group, University of California, San Francisco

[2] Department of Pharmaceutical Chemistry, University of California, San Francisco

[3] Institute of Molecular Systems Biology, ETH Zürich

[4] Department of Microbiology and Immunology, University of California, San Francisco

[5] Department of Bioengineering and Therapeutic Sciences, University of California, San Francisco

Introduction

Within a tumor, extensive genetic variation between cells can lead to phenotypic heterogeneity, complicating clinical decision-making in cancer medicine (Burrell *et al*., 2013) and enabling tumors to become resistant to drug treatment via selection of pre-existing clones (Hata *et al*., 2016). Beyond genetic diversity, recent works have suggested that phenotypic heterogeneity can arise from non-genetic mechanisms (Brock *et al*., 2009) such as epigenetic markers (Feinberg *et al*., 2016; Flavahan *et al*., 2017; Hinohara *et al*., 2018) or metabolism (Hangauer *et al*., 2017; Roesch *et al*., 2013; Viswanathan *et al*., 2017). For acute myeloid leukemias (AML) which have relatively low genetic diversity (Tian *et al*., 2015), several of the most recurrently mutated genes encode epigenetic modifiers (Cancer Genome Atlas Research Network, 2013), suggesting that AML fitness grows with "epigenetic diversity" (Li *et al*., 2016). Consistently, mutations in epigenetic modifiers are frequently found to persist through remission to relapse in AMLs (Ding *et al*., 2012; Corces-Zimmerman *et al*., 2014; Rothenberg-Thurley *et al*., 2018).

We have previously shown that cells with a loss of the epigenetic modifier TET2 have increased survival in the presence of chemotherapy due in part to altered cell state switching dynamics (Morinishi *et al*., 2020; Chapter 2). In the course of our study, we also observed a change in growth rate and metabolic activity in isogenic TET2 mutants compared to the WT cell line. These observations suggest a link between epigenetics and metabolism and present an interesting possibility: if TET2 mutants have a distinct metabolic state, they may also have distinct metabolic vulnerabilities. Indeed, recent works by Hangauer *et al*. and Viswanathan *et al*. have shown that drug-tolerant cancer cells can be particularly vulnerable to inhibition of the lipid hydroperoxidase GPX4 (Hangauer *et al*., 2017; Viswanathan *et al*., 2017). Here we describe the metabolic effects of TET2 loss in AML cells, and study the effects of metabolic perturbation on AML cell chemosurvival *in vitro*.

Results and Discussion

To better understand how TET2 mutation changes cellular metabolism, we measured mitochondrial function in the isogenic KG1 cell lines with the Seahorse XFe96 Analyzer (Agilent). Respiration in $TET2^{KO}$ cells was significantly higher than WT cells as measured by the oxygen consumption rate (OCR) for maximal respiration and spare respiratory capacity (Figure 3.1). KG1 $TET2^{KO}$ cells showing increased respiration is consistent with the cell line's increased cell proliferation (Figure B.2).

We also profiled the metabolomics of our isogenic KG1 and Thp1 cell lines with flow-injection time-of-flight mass spectrometry (FIA TOF-MS). For 1199 putatively annotated ions, normalized abundances measured from 8 replicate measurements showed high reproducibility between samples (median CV is 11% for non-washed samples and 16% for washed samples). Cells washed with wash buffer prior to quenching (see Methods) had slightly different fold-change values compared to unwashed cells (Figure 3.2 and Figure 3.3), but the overall trends were not affected (Figure 3.4). TET2 mutation appeared to have a stronger effect on the metabolome in KG1 cells than in Thp1 (Figure 3.2 and Figure 3.3, with ~100/1199 metabolites significantly affected), with few metabolites showing a fold-change in the same direction across cell lines (Figure 3.4). This is consistent with our RNAseq and DNA methylation profiling, which showed a stronger effect in KG1 TET2 mutants compared to Thp1 (Figure A.3). These differences may be due to the disparate methods of knocking out/down TET2 (Cas9 editing in KG1, shRNA in Thp1) or differences in cell type (KG1: $CD34^{+}CD38^{+}$ myeloblast; Thp1: $CD34^{-}$ macrophage myeloblast).

To better understand which metabolic pathways were affected by TET2 mutation, we performed pathway analysis for significantly altered metabolites with the MetaboAnalyst tool (Chong *et al*., 2019) using "KEGG (Previous)" pathways. KG1 TET2 mutant cells were significantly enriched for TCA cycle metabolites (L-Malic acid, cis-Aconitic acid, Citric acid, and Fumaric acid; Tables 3.1 and 3.2) and glutathione metabolism compared to $TET2^{WT}$, regardless of washing protocol. Moreover, ATP and several co-factors (e.g. Acetyl-CoA) were elevated in KG1 TET2 mutants, consistent with the observed increase in respiration in the KG1 $TET2^{KO}$ cell line (Figure 3.1).

Given the differences in metabolic activity in KG1 TET2^{KO} cells, we wondered whether modulation of metabolic activity could affect drug survival. In Chapter 2, we showed that TET2^{KO} AML cells (which are often found to persist through remission to relapse in patients) have increased survival in the presence of chemotherapy (1μM AraC) due in part to altered cell state switching dynamics. Further, we showed that slowing switching dynamics with extracellular effectors altered drug survival (Figure 2.4 and Figure A.12). Effectors or signaling molecules that preferentially suppress the survival of TET2^{KO} cells in drug may help us identify key metabolic vulnerabilities in the more drug-resistant TET2^{KO} cells, which could then be leveraged to reduce the rate of relapse in patients with TET2-mutant AMLs. To identify effectors, we expanded our panel to 35 drugs or biologics spanning glycolysis, autophagy, and oxidative phosphorylation (Table C.1), and generated 72h dose-response curves for a 2X dilution series of each effector in both KG1 TET2^{WT} and TET2^{KO} cell lines treated with 0 or 1μM AraC (Figure C.1).

Interestingly, our tested effectors appeared to have one of four types of dose-response curves (Figure 3.5): Effectors that at the given concentrations had no discernable effect on survival either in or out of drug (**Group 1:** 2HG, DCA, DNP, Fructose, IU1, Kynurenine, Mannose, Metyrapone, MRT68921, MS, NAC, Na-Lactate, Na-Pyruvate, Oxamate, PTIO, Sulfasalazine); Effectors with strong effects that were largely the same in and out of drug (**Group 2:** TSA, BrPyr, DHA); Effectors that enabled better Tet2^{KO} growth out of drug, but did not notably alter drug survival (**Group 3:** ATRA, Bafilomycin, BSO, CB-839, Chloroquine, Lopermaide, Metformin, Rotenone, Valproic acid); and Effectors that showed higher Tet2^{WT} survival compared to Tet2^{KO} in drug (**Group 4:** 2DG, CCCP, DS, EA, IFNg, Torin).

The observed effect of disulfiram and IFNg on TET2^{KO} cell survival in 1μM AraC is consistent with previous measurements from Chapter 2 (Figure 2.4 and Figure A.12). Unfortunately, no single mechanism uniformly explains the change in drug survival observed for all Group 4 effectors. Looking at each effector individually, however, offers some insight. For example, disulfiram inhibits the enzyme acetaldehyde dehydrogenase and thus may slow production of acetyl-CoA, a significantly enriched metabolite in KG1 TET2^{KO} cells (Figure 3.2). Further work will be required to understand the mechanisms that cause effectors in Group 4 to differentially affect TET2 mutants.

Together, these data provide strong evidence that KG1 TET2^{KO} cell lines have significantly altered metabolic activity compared to TET2^{WT}. However, future work will be required to more precisely describe the metabolic differences and, potentially, identify compounds or treatment strategies that preferentially target the more drug-tolerant TET2-mutant AML cells.

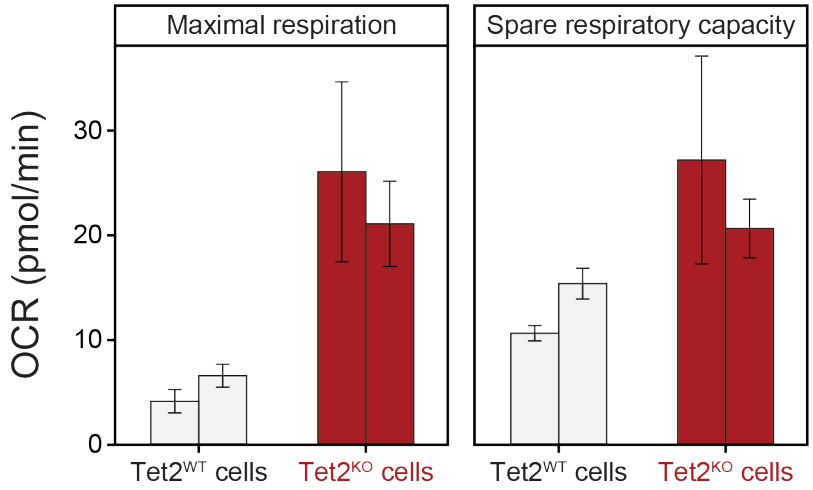

Figure 3.1: KG1 TET2^{KO} cells have higher levels of respiration than TET2^{WT}
Oxygen consumption rate (OCR) was measured for four technical replicate wells of two replicates per cell line (KG1 TET2^{WT} or TET2^{KO}) on the Seahorse XFe96 (Agilent). Shown are values for maximal respiration and spare respiratory capacity. Low OCR values here may be a result of few suspension cells remaining in the plate after preparatory wash steps.

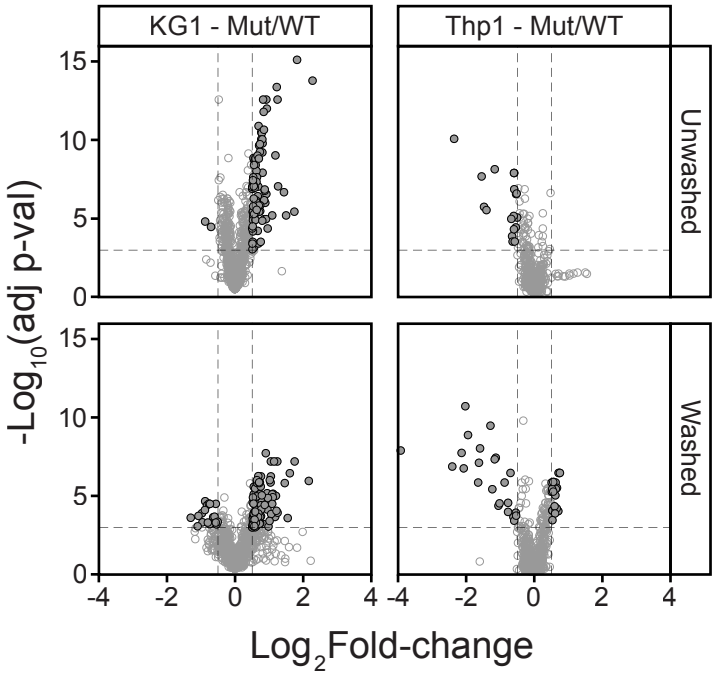

Figure 3.2: Volcano plot of KG1 and Thp1 cell metabolomics
Each circle represents an ion, data shown is the mean of 8 replicate measurements. Filled circles: metabolites with absolute $Log_2FC > 0.5$ and adjusted p-value < 0.001.

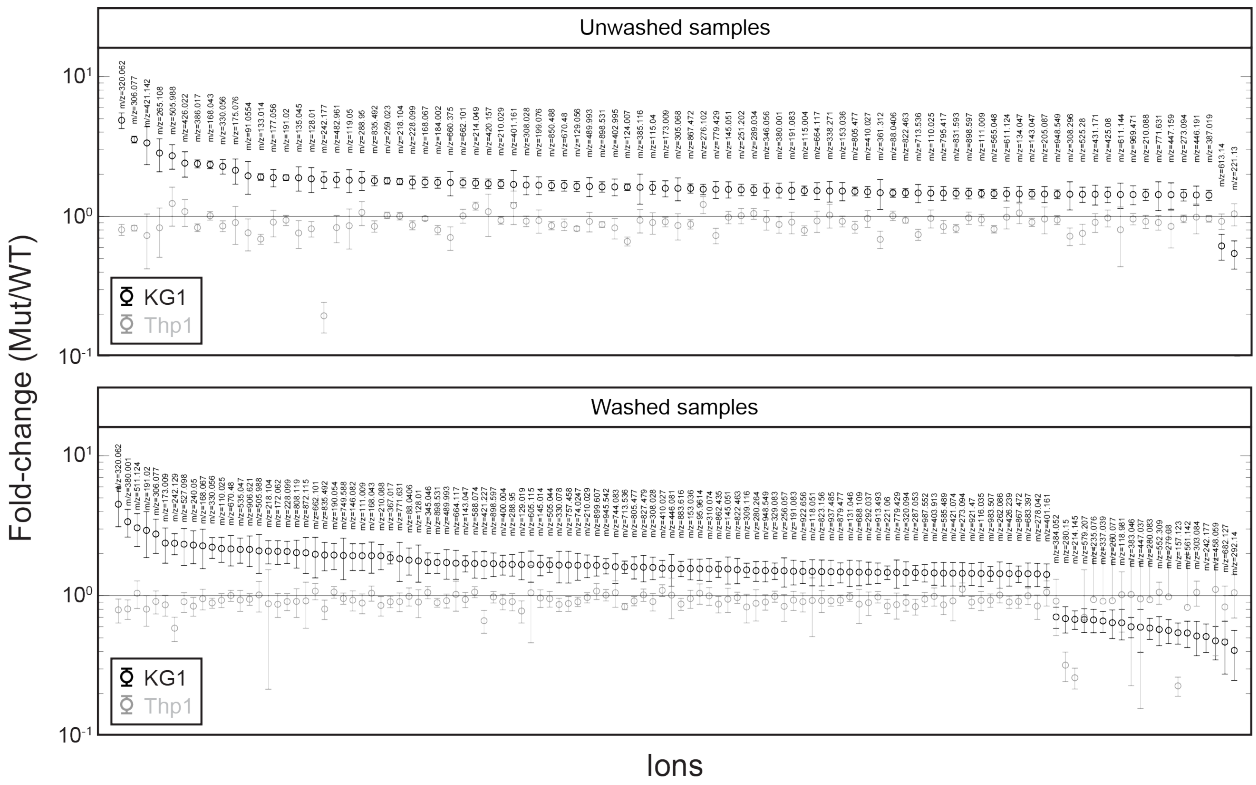

Figure 3.3: Sorted significantly altered features in TET2 mutants
Shown are features with absolute $Log_2FC > 0.5$ and adjusted p-value < 0.001 for KG1 (black) and Thp1 (gray). Labels correspond to mass-to-charge ratios (m/z).

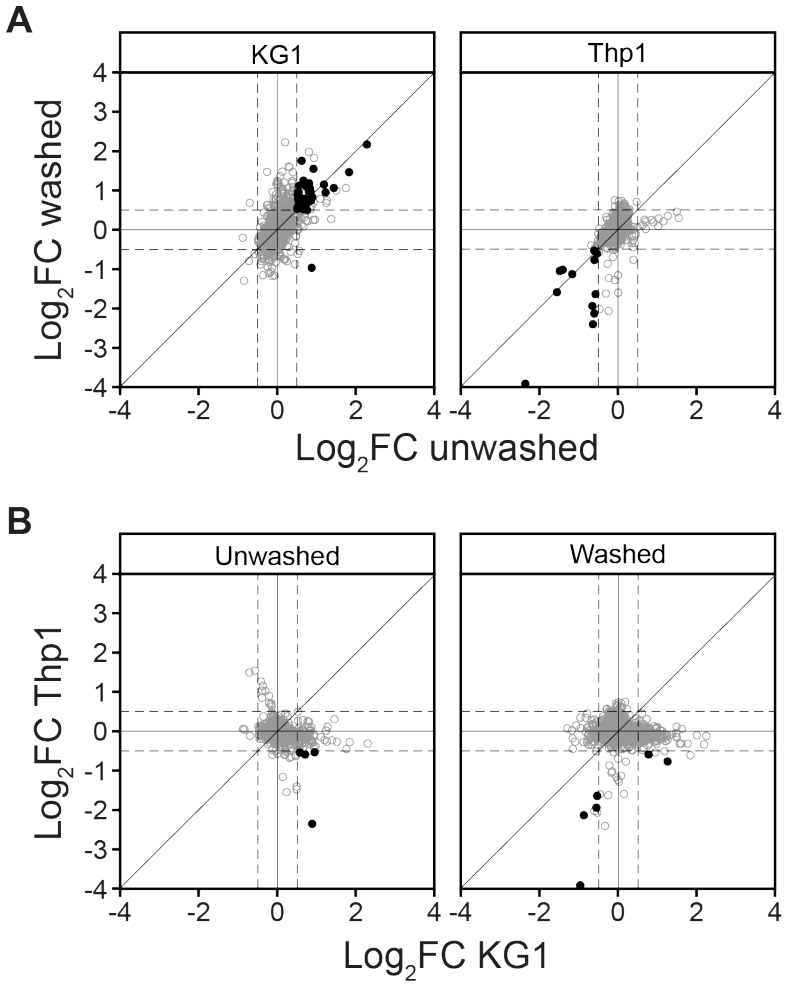

Figure 3.4: Fold-change per features for KG1 and Thp1 cells with or without washing
(A) For each cell line, a comparison of feature fold-change in TET2 mutants for washed and unwashed samples is shown. (B) For either washed or unwashed samples, a comparison of feature fold-change in TET2 mutants for the KG1 and Thp1 cell lines is shown. Filled circles: metabolites with absolute $Log_2FC > 0.5$ and adjusted p-value < 0.001.

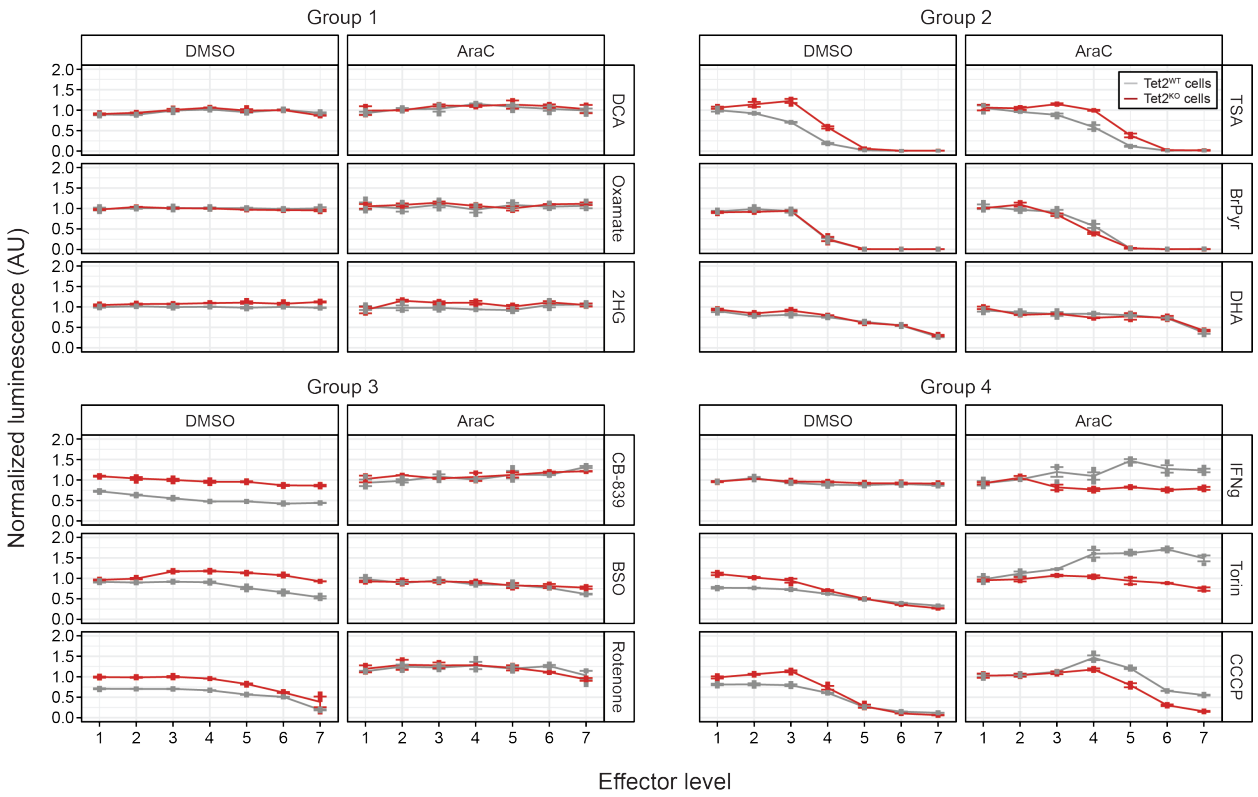

Figure 3.5: Effector panel identifies modulators of KG1 chemosensitivity
KG1 TET2^{WT} and TET2^{KO} cells were treated with a 2X dilution series of a panel of effectors (Table C.1, where Effector Level 7 is the highest dose) and either DMSO or 1μM AraC for 72h prior to cell viability measurement with CellTiter-Glo. Effectors appeared to have one of four types of dose-response curves: (1) no discernable effect on survival at the given concentrations, (2) strong effect that is very similar in and out of drug, (3) some effect on TET2^{KO} growth out of drug, but no notable change in drug survival, and (4) higher TET2^{WT} cell survival compared to TET2^{KO} in drug. Shown are dose-response curves for 3 representative effectors from each group (mean ± s.e., n=3).

Table 3.1: Pathway analysis of unwashed KG1 samples
KEGG pathways enriched in significantly altered putatively annotated ions in unwashed KG1 TET2^{KO} cells.

Pathway	Total	Expected	Hits	p-value	-log(p)	Holm adj	FDR	Impact
Citrate cycle	20	0.35729	4	3.48e-4	7.9638	0.02783	0.02782	0.17883
Pantothenate and CoA biosynthesis	27	0.48234	4	1.15e-3	6.7678	0.09087	0.04601	0.32719
Pyrimidine metabolism	60	1.0719	5	3.783e-3	5.5773	0.29507	0.10088	0.05288
Purine metabolism	92	1.6435	6	5.148e-3	5.2692	0.39636	0.10295	0.11185

Table 3.2: Pathway analysis of washed KG1 samples
KEGG pathways enriched in significantly altered putatively annotated ions in washed KG1 TET2^{KO} cells.

Pathway	Total	Expected	Hits	p-value	-log(p)	Holm adj	FDR	Impact
beta-Alanine metabolism	28	0.46531	5	7.28e-5	9.5275	0.00583	0.00583	0.31334
Citrate cycle	20	0.33236	4	2.62e-4	8.2481	0.02068	0.01047	0.23915
Glutathione metabolism	38	0.63149	4	3.21e-3	5.7405	0.25062	0.08568	0.3244
Glyoxylate and dicarboxylate metab.	50	0.83091	4	8.68e-3	4.7462	0.6687	0.15098	0.00881
Pantothenate and CoA biosynthesis	27	0.44869	3	9.44e-3	4.6632	0.71714	0.15098	0.29143

Methods

Cell lines, reagents, and cell culture

KG1 cells and Thp1/Thp1 TET2^{KD} cells were generous gifts from the Shannon lab at UCSF and Levine lab at MSKCC, respectively. Cells were cultured in ATCC-recommended media and incubated at 37°C with 5% CO_2. All cell lines were maintained in 75ml or 250ml Suspension Culture flasks (CellTreat) between 0.5e6 and 1e6 cells/ml.

To make TET2^{KO} cells in the KG1 cell line, guides targeting exon 3 of TET2 (an exon with frequent indel mutations in patients with MDS; Smith *et al*., 2010) were designed in Benchling (5'-CACCGAGGCCAATTAAGGTGGAACC-3'). pSpCas9(BB)-2A-Puro (PX459) V2.0 was a gift from Feng Zhang (Addgene plasmid #62988). Guides were cloned into PX459 using BbsI sites, and vectors transfected into KG1 cells with the Cell Line Nucleofector Kit R (Lonza Group, Basel, Switzerland) with protocol V-001 and according to the manufacturer's protocol. After 48h, FITC-positive cells were isolated via FACS and expanded for 10 days. The TET2 mutation created a frameshift mutation causing truncation (Figure A.1) upstream of the conserved catalytic domain of TET2, consistent with tumors observed in patients (Weissman *et al*., 2012).

Seahorse

Cell-Tak solution (Agilent) was adsorbed to an XF96 Cell Culture Microplate per manufacturer instructions. Four wells of two technical replicates each for KG1 TET2^{WT} and TET2^{KO} cells were loaded into the appropriate wells and centrifuged prior to assay administration and measurement on the XFe96 (per manufacturer's instructions, Agilent).

Metabolomics

Two technical replicates of 5e5 cells for each cell line were pelleted at 300g for 3 minutes. Half of the samples were washed once with wash buffer (75mM Ammonium bicarbonate, pH 7.4, 37°C). All samples were then resuspended in 500µL quenching solution (40:40:20 CAN:MeOH:H_2O, -20°C) and vortexed prior to incubation at -20°C for 1 hour. Samples were centrifuged at 300g for 3 minutes, and 4x50µL of each supernatant was transferred into conical storage plates as technical replicates. Plates were

heat-sealed and stored at -80°C. Samples were analyzed by double injection on an iFunnel 6650 Q-TOF (Agilent). A custom script in MATLAB was used to calculate the mean and standard deviation of fold-change intensity for each WT/KO cell line pair (for washed or unwashed samples), the false discovery rate, and the adjusted p-value (t-test; Storey, 2002).

Cell viability assays

For cell viability assays, cells were plated in 384-well plates at 1e5 cells/ml either with a Matrix WellMate (three replicate plates; Thermo Scientific) or by hand (one replicate plate) and treated with drug and/or effector for 72h at 37°C with 5% CO_2. Effectors and drug were either added with the Echo (Labcyte, San Jose, CA) or by hand. Plates were allowed to cool to room temperature before combining well-mixed cells 1:1 with room temperature CellTiter-Glo 2.0 (Promega) in opaque, white tissue culture plates (Corning). Reactions were allowed to proceed according to the manufacturer's protocol, and luminescence was read out with the Biotek H4 plate reader (BioTek) in the Center for Advanced Technology at UC San Francisco. Luminescence in each plate was normalized to the median signal from all "no effector" wells (either only DMSO or only AraC, as relevant) prior to visualization. Results shown are representative of two independent experiments.

References

Brock A, Chang H & Huang S. (2009). Non-genetic heterogeneity a mutation-independent driving force for the somatic evolution of tumours. Nature Reviews Genetics, 10(5), 336–342.

Burrell RA, McGranahan N, Bartek J & Swanton C. (2013). The causes and consequences of genetic heterogeneity in cancer evolution. Nature, 501(7467), 338–345.

Cancer Genome Atlas Research Network *et al*. (2013). Genomic and epigenomic landscapes of adult de novo acute myeloid leukemia. N. Engl. J. Med. 368, 2059–2074.

Chong J, Wishart DS & Xia J. (2019). Using MetaboAnalyst 4.0 for Comprehensive and Integrative Metabolomics Data Analysis. Current Protocols in Bioinformatics, 68(1), 1–128.

Corces-Zimmerman MR, Hong WJ, Weissman IL, Medeiros BC, Majeti R. (2014). Preleukemic mutations in human acute myeloid leukemia affect epigenetic regulators and persist in remission. Proceedings of the National Academy of Sciences, 111(7), 2548–2553.

Ding L. *et al*. (2012). Clonal evolution in relapsed acute myeloid leukaemia revealed by whole-genome sequencing. Nature 481, 506–510.

Feinberg AP, Koldobskiy MA, Göndör A. (2016). Epigenetic modulators, modifiers and mediators in cancer aetiology and progression. Nature Reviews Genetics, 17, 284-299.

Flavahan WA, Gaskell E, Bernstein BE. (2017). Epigenetic plasticity and the hallmarks of cancer. Science, 357(6348).

Hata AN, Niederst MJ, Archibald HL, Gomez-Caraballo M, Siddiqui FM, Mulvey HE, Maruvka YE, Ji F, Bhang HC, Krishnamurthy Radhakrishna V, Siravegna G, *et al*. (2016). Tumor cells can follow distinct evolutionary paths to become resistant to epidermal growth factor receptor inhibition. Nature Medicine, 22(3), 262–269.

Hangauer MJ, Viswanathan VS, Ryan MJ, Bole D, Eaton JK, Matov A, Galeas J, Dhruv HD, Berens ME, Schreiber SL, *et al*. (2017). Drug-tolerant persister cancer cells are vulnerable to GPX4 inhibition. Nature, 551:247-250.

Hinohara K, Wu HJ, Vigneau S, McDonald TO, Igarashi KJ, Yamamoto KN, Madsen T, Fassl A, Egri SB, Papanastasiou M, *et al*. (2018). KDM5 histone demethylase activity links cellular transcriptomic heterogeneity to therapeutic resistance. Cancer Cell, 34(6), 939-953.e9.

Li S, Mason CE & Melnick A. (2016). Genetic and epigenetic heterogeneity in acute myeloid leukemia. Curr. Opin. Genet. Dev. 36, 100–106.

Roesch A, Vultur A, Bogeski I, Wang H, Zimmermann KM, Speicher D, Körbel C, Laschke MW, *et al*. (2013). Overcoming intrinsic multidrug resistance in melanoma by blocking the mitochondrial respiratory chain of slow-cycling JARID1B(high) cells. Cancer Cell, 23:811–25.

Rothenberg-Thurley M, Amler S, Goerlich D, Köhnke T, Konstandin NP, Schneider S, Sauerland MC, Herold T, Hubmann M, Ksienzyk B, *et al*. (2018). Persistence of pre-leukemic clones during first remission and risk of relapse in acute myeloid leukemia. Leukemia, 32(7), 1598–1608.

Smith AE, Mohamedali AM, Kulasekararaj A, Lim Z, Gäken J, Lea NC, Przychodzen B, Mian SA, Nasser EE, Shooter C, *et al*. (2010). Next-generation sequencing of the TET2 gene in 355 MDS and CMML patients reveals low-abundance mutant clones with early origins, but indicates no definite prognostic value. Blood, 116(19), 3923–3932.

Storey JD. (2002). A direct approach to false discovery rates. Journal of the Royal Statistical Society: Series B (Statistical Methodology), 64(3), 479–498.

Tian R, Basu MK & Capriotti E. (2015). Computational methods and resources for the interpretation of genomic variants in cancer. BMC Genomics 16 Suppl 8, S7.

Viswanathan VS, Ryan MJ, Dhruv HD, Gill S, Eichhoff OM, Seashore-Ludlow B, Kaffenberger SD, Eaton JK, Shimada K, Aguirre AJ, *et al*. (2017). Dependency of a therapy-resistant state of cancer cells on a lipid peroxidase pathway. Nature, 547:453–457.

Weissmann S, Alpermann T, Grossmann V, Kowarsch A, Nadarajah N, Eder C, Dicker F, Fasan A, Haferlach C, Haferlach T, Kern W, Schnittger S & Kohlmann A. (2012). Landscape of TET2 mutations in acute myeloid leukemia. Leukemia, 26(5), 934–942.

Appendix A

Appendix A: Supplemental Material for Chapter 2

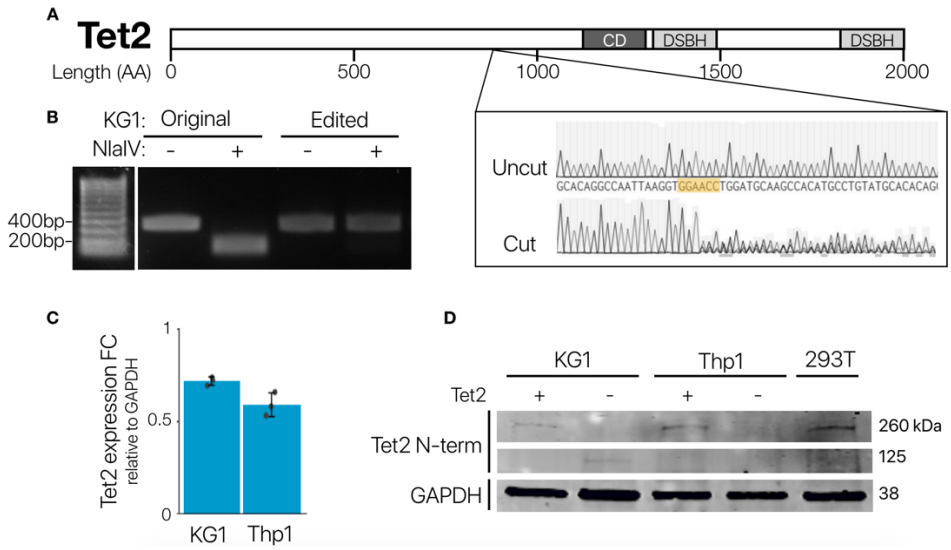

Figure A.1: Confirming loss of Tet2 expression
(A) An outline of the gene TET2 showing sequence the final gRNA targeted, and a bulk Sanger sequencing trace of the locus before and after plasmid nucleofection. (B) Upon editing, a NlaIV cut site is removed. PCR amplicons of a 400bp region around the target locus were run on an agarose gel with or without NlaIV. (C) TET2 expression fold-change, as measured by RT-qPCR in TET2KO pool relative to the parental cell line, shows decrease in expression relative to GAPDH fold-change. (D) Western blot of both cell lines before and after TET2 knockout with TET2 N-terminus targeting antibody. 293T is a positive control. TET2 band is lost at 260kDa.

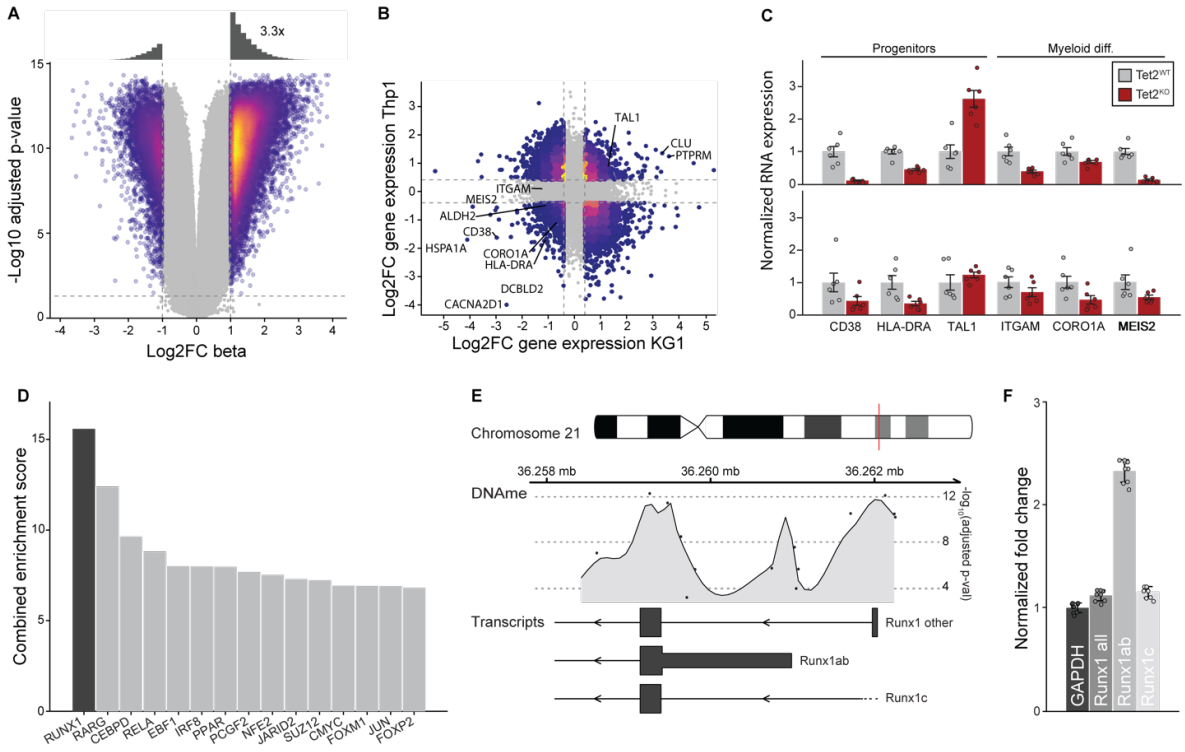

Figure A.2: Changes in DNA methylation and RNA expression in TET2^{KO} cells
(A) Representative volcano plot of log_2fold-change in DNA methylation beta values after TET2 loss (shown for KG1). Each point is one CpG. As expected, TET2^{KO} cells are more hypermethylated at CpGs across the genome. (B) Log_2fold-change of RNA expression counts from RNAseq after TET2 loss for KG1 (x-axis) and Thp1 (y-axis). Each point represents a gene, and some similarly differentially expressed genes are labeled. Dotted lines denote a threshold of 1.5X change. (C) Genes that modulate differentiation and stemness are differentially expressed as measured by RNAseq after TET2^{KO} in both KG1 and Thp1 (Figure A.2, mean ± s.e., n=6). TET2^{KO} cell lines show decreased expression of myeloid commitment markers (ITGAM, CORO1A) and increased expression of markers associated with stemness (CD38^{-}, HLA-DRA^{-}, TAL1) compared to TET2^{WT}. (D) Genes that are differentially expressed after TET2 loss are enriched for targets of Runx1. Results from the "ChEA 2016" enrichment analysis of significantly differentially expressed genes after TET2^{KO}, showing transcription factors in order of "Combined Enrichment Score". (E) Analysis of DNA methylation data found that the most significant DMR in TET2^{KO} cells was the proximal promoter of Runx1. An ideogram representing the chromosomal location of Runx1 (top), the significance of differential methylation in the region (middle), and relevant transcripts (bottom) are shown. (F) Runx1a isoforms are differentially expressed in TET2^{KO} cells. Fold-change of Runx1 isoforms in KG1 TET2^{KO} cells is shown, normalized to GAPDH fold-change measured by RTqPCR (mean ± s.e., n=8).

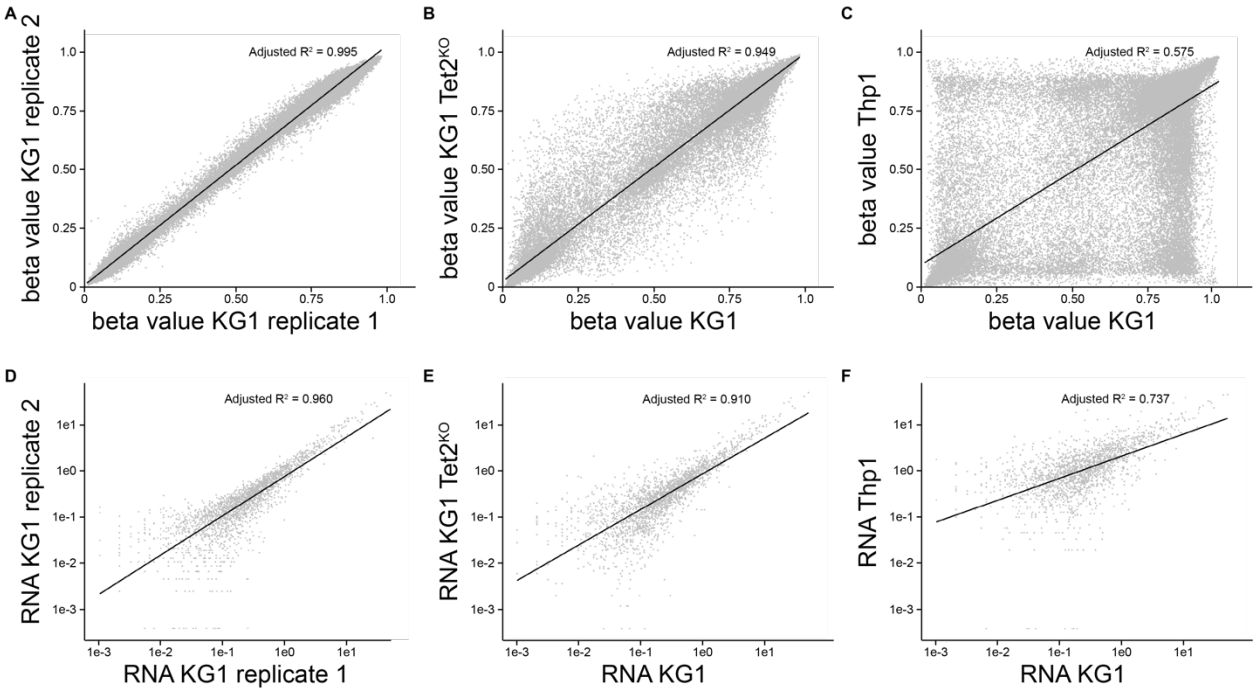

Figure A.3: Correlation of DNA methylation and RNA expression across replicates, treatment conditions, and cell lines

Representative DNA methylation (A-C) and RNA expression (D-F) correlation between (A,D) technical replicates, (B,E) $TET2^{WT}$ and $TET2^{KO}$, and (C,F) different untreated cell lines.

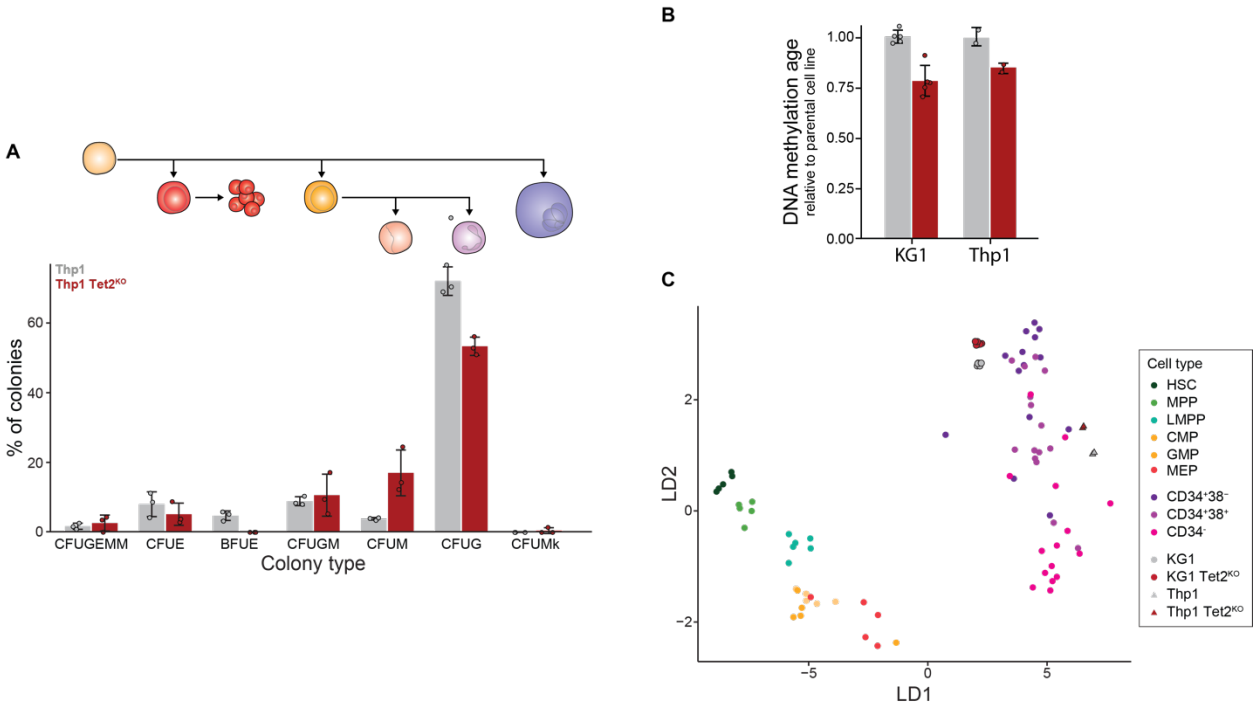

Figure A.4: TET2^{KO} myeloblasts are more stem-like than their TET2^{WT} counterparts
(A) Thp1 TET2^{KO} cell lines produce a lower percentage of colonies associated with specialized granulocytic cells (CFUG) compared to TET2^{WT} cells (mean ± s.e., n=3) in methylcellulose colony forming assays. (B) DNA methylation age as measured by the Horvath calculator for KG1 and Thp1 cells for TET2^{WT} (gray) and TET2^{KO} (red) (mean ± s.e., n=3). (C) Low-dimensional representation of DNA methylation profiles after using LDA to find the space that best separates HSCs, CD34^{+}38^{-}, and CD34^{-} cells in a reference dataset (Jung *et al*., 2015). Each point is one sample, and the color corresponds to the cell type of the sample. The two LD axes are readily interpretable: LD1 separates normal hematopoietic cells from leukemic and increasing LD2 moves up the known hematopoietic hierarchies. KG1 and Thp1 TET2^{KO} cells have greater LD2 values than their corresponding TET2^{WT} counterparts.

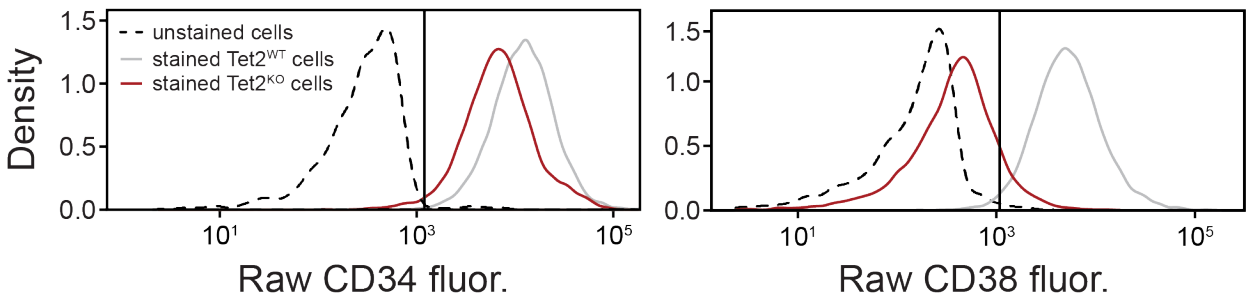

Figure A.5: Unstained cell populations used to define CD38^{lo} and CD38^{hi} cell states
Shown are representative distributions of raw CD34 (left panel) and CD38 (right panel) fluorescence for both stained and unstained KG1 cells. To define "lo" and "hi" states, the fluorescence of unstained WT cells were measured; "lo" or "hi": cell with intensities below or above (resp.) 99.5% (vertical line) of the distribution.

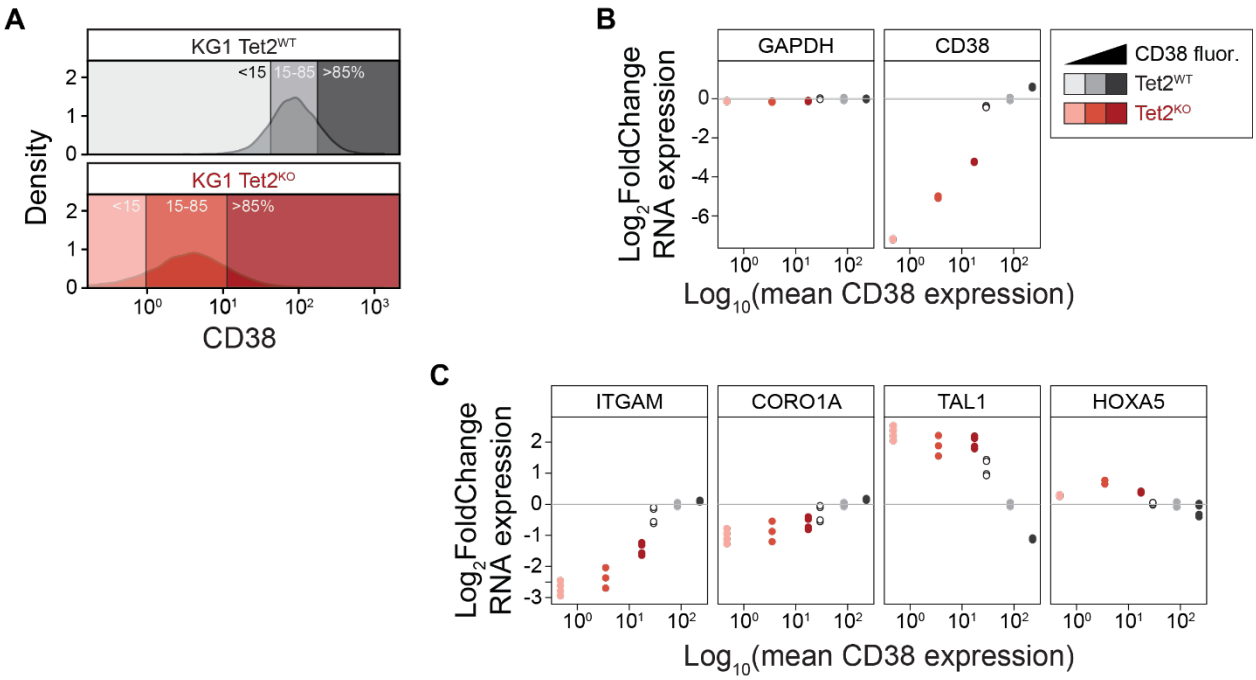

Figure A.6: Expression of stem-like or differentiated genes correlate with CD38 expression
(A) KG1 TET2^{WT} and TET2^{KO} cells were fractionated by CD38 expression by FACS (<15th percentile, 15th-85th percentile, >85th percentile) prior to RNA extraction and quantification by RTqPCR. Log_2fold-change expression compared to GAPDH was calculated for each gene and plotted as a function of the mean CD38 expression (fluorescence) of each sorted population (as measured by flow cytometry). (B) Shown are the Log_2fold-change RNA expression of GAPDH (reference gene control) and CD38 as a function of mean CD38 fluorescence. CD38 RNA expression increases with increased CD38 surface marker expression, while GAPDH levels remain constant. (C) Shown are the Log_2fold-change RNA expression of myeloid differentiation markers ITGAM and CORO1A, and progenitor markers TAL1 and HOXA5 compared to GAPDH.

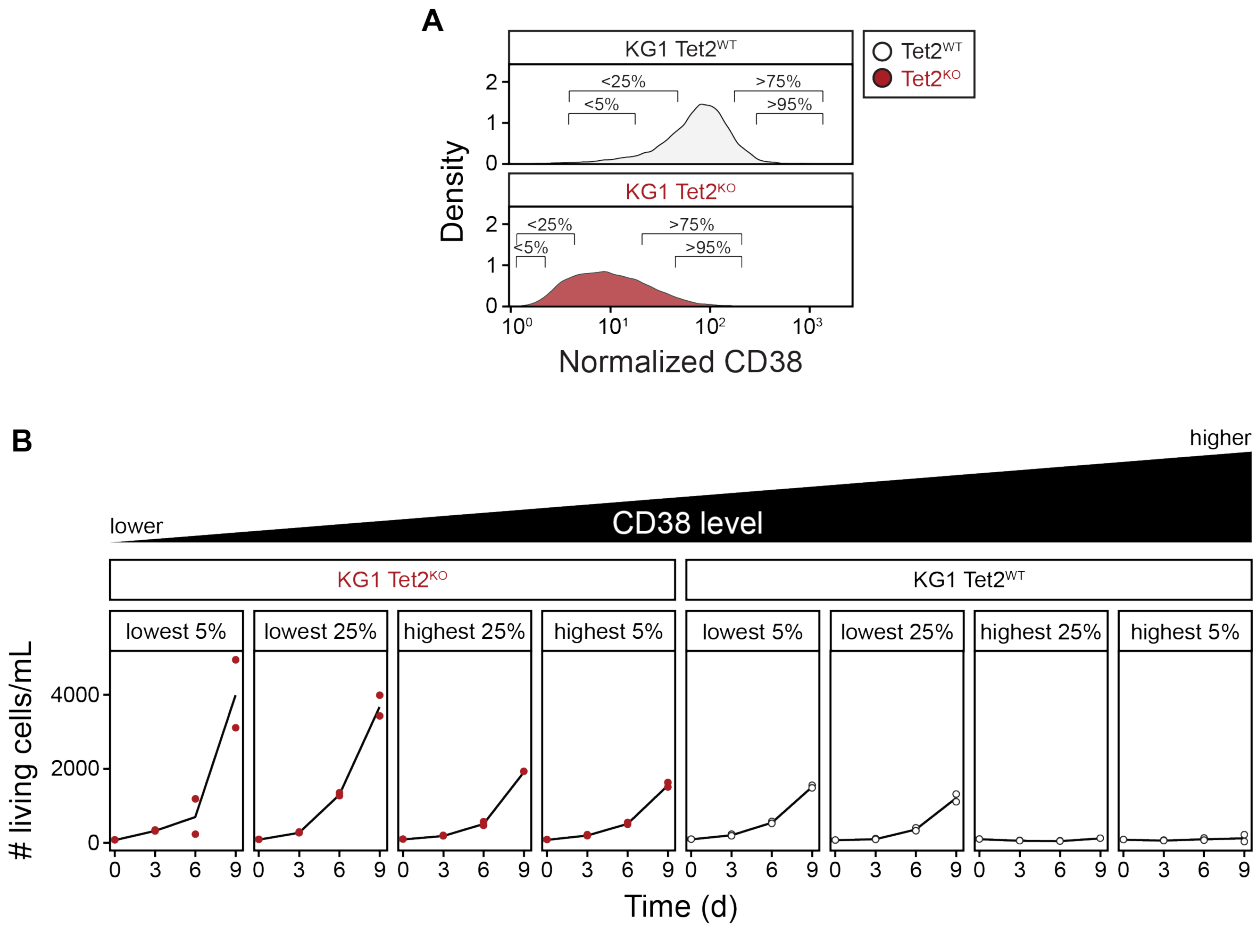

Figure A.7: Cells with high CD38 expression have a lower growth rate
(A) KG1 TET2^{WT} and TET2^{KO} cells were partitioned by CD38 expression (<5th percentile, <25th percentile, >75th percentile, >95th percentile), and monitored for 9 days. (B) At each timepoint, cells were resuspended in the same volume of buffer and assessed by flow cytometry at a constant flow rate. Thus, the number of living cells per unit volume observed during flow can be used as a proxy for living cell density. The lowest CD38-expressing populations (far left) appear to have the highest growth rate. The highest CD38-expressing subpopulations (far right) have the lowest growth rate, resulting in a flat trend.

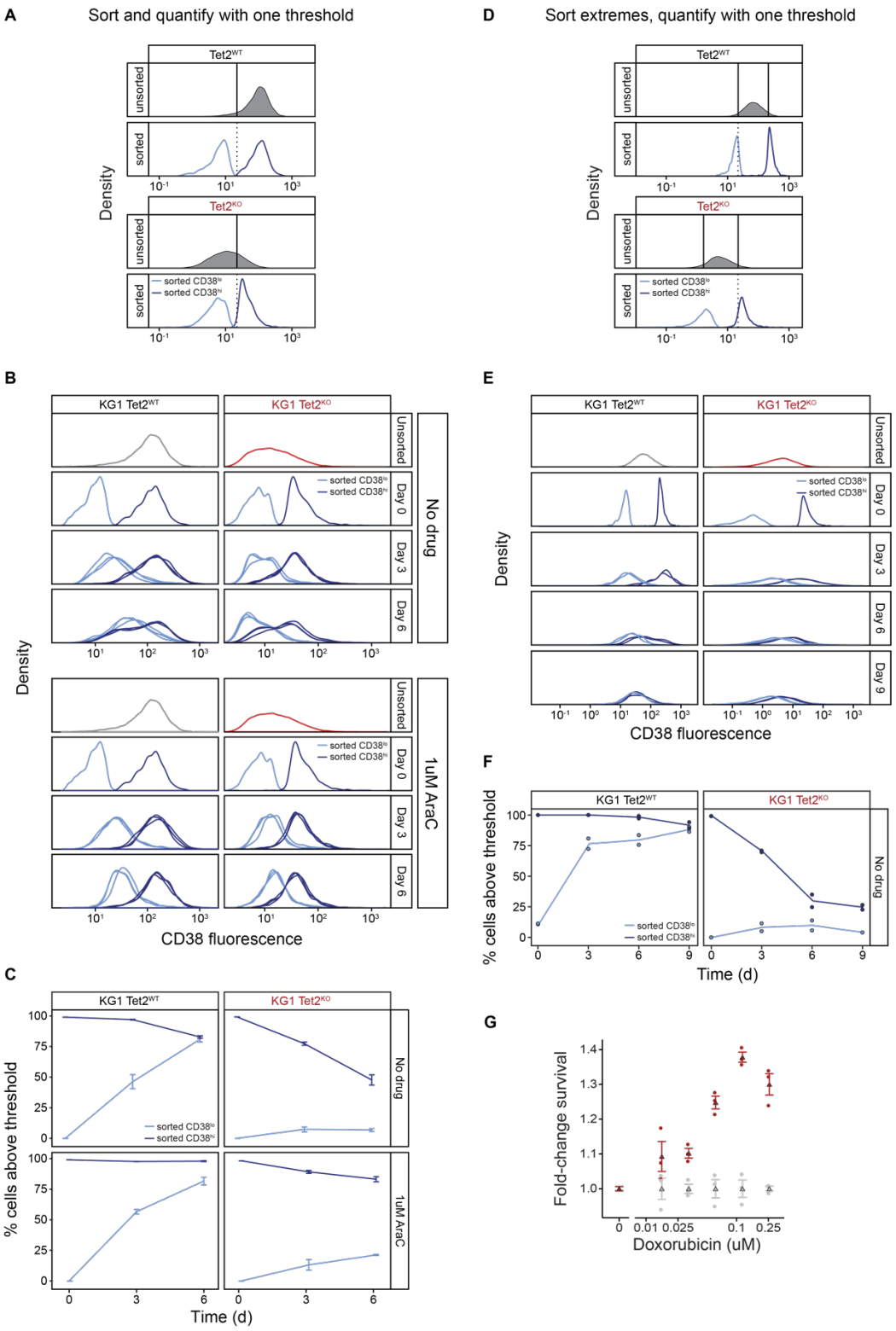

Figure A.8: TET2 status and drug treatment alter the dynamics of transitions between CD38-defined cell states

(A-C) Results of sorting CD38 based on single threshold. (A) Schematic showing how KG1 and KG1 TET2^{KO} cells (gray) were sorted based on CD38 expression (light blue: CD38 "low", dark blue: CD38

"high"; solid vertical line: flow cytometry sort gates; dotted vertical line: the threshold was used to call cells as CD38 "lo" or "hi"; see Figure A.5). (B) Shown are the distributions of CD38 expression measured by flow cytometry over the course of 6 days in 0μM or 1μM AraC, colored by the original sorted population (n=3). (C) The percent of cells above the quantification threshold (vertical dashed line in A) are shown for all sorted populations over time (mean ± s.e.). (D-F) Results of sorting CD38 from extremes of distributions. (D) the "tails" (bottom/top 5%) of the CD38 distributions of KG1 $TET2^{WT}$ and $TET2^{KO}$ cells were isolated by FACS, and their CD38 expression was monitored for 9 days (light blue: low, dark blue: high; solid vertical line: flow cytometry sort gates; dotted vertical line: the same threshold from (A) used to call cells as CD38 "lo" or "hi"). (E) Shown are the CD38 densities of the sorted populations from (D) over the course of 9 days, colored by the original sorted population (n=2). (F) The percent of living cells above the quantification threshold (vertical dashed line in D) relaxes to a steady-state over time. The highest 5% of CD38-expressing KG1 $TET2^{WT}$ cells (theoretically the most differentiated cells) still appear to return to steady state over time, but the number of living cells in these samples is dramatically lower than the rest (see Figure A.7). (G) Fold-change survival of KG-1 $TET2^{KO}$ cells (red) relative to WT (gray) in varying concentrations of Doxorubicin at 72 hrs.

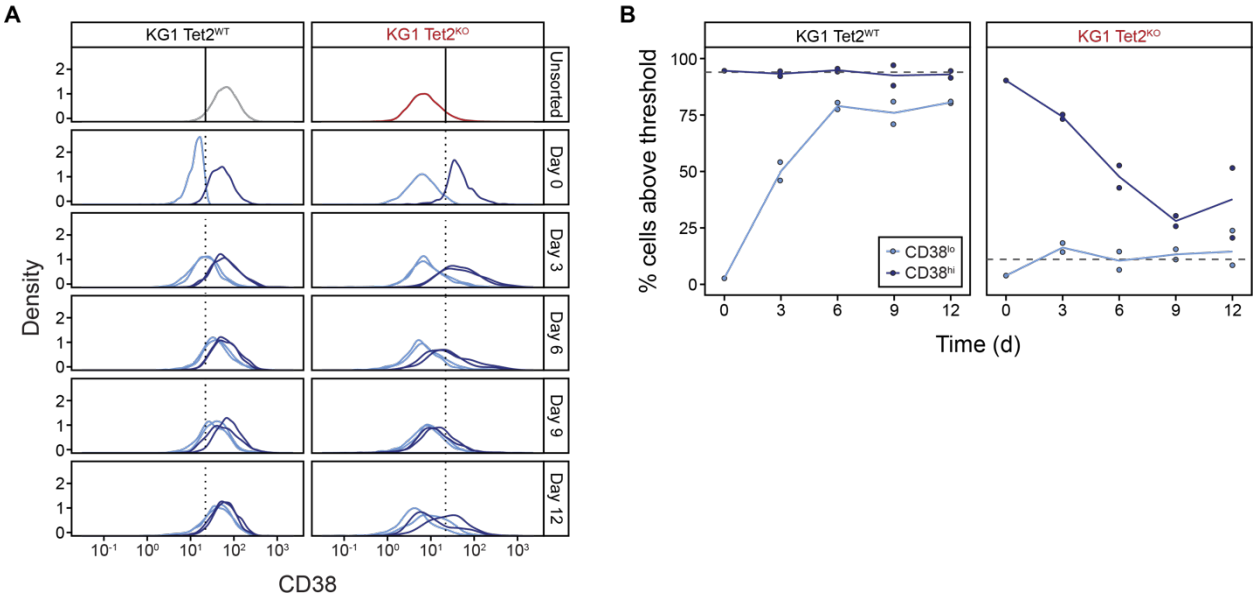

Figure A.9: Sorted KG1 TET2^{WT} and TET2^{KO} cells return to steady-state
(A) KG1 TET2^{WT} (gray) and TET2^{KO} (red) cell populations (top row) were sorted based on CD38 expression (light blue: CD38 "lo", dark blue: CD38 "hi"; solid vertical line: flow cytometry sort gates; dotted vertical line: the threshold used to call cells as CD38 "lo" or "hi"; see Figure A.5), and their CD38 surface marker expression was monitored for 12 days (n=2). Shown are the distributions of CD38 expression before and after cell sorting. (B) The percent of living cells above the threshold in (A) is shown. The expected steady-state (horizontal dashed line) for each cell line is the percent of living cells above the threshold in (A) in unsorted populations on day 0.

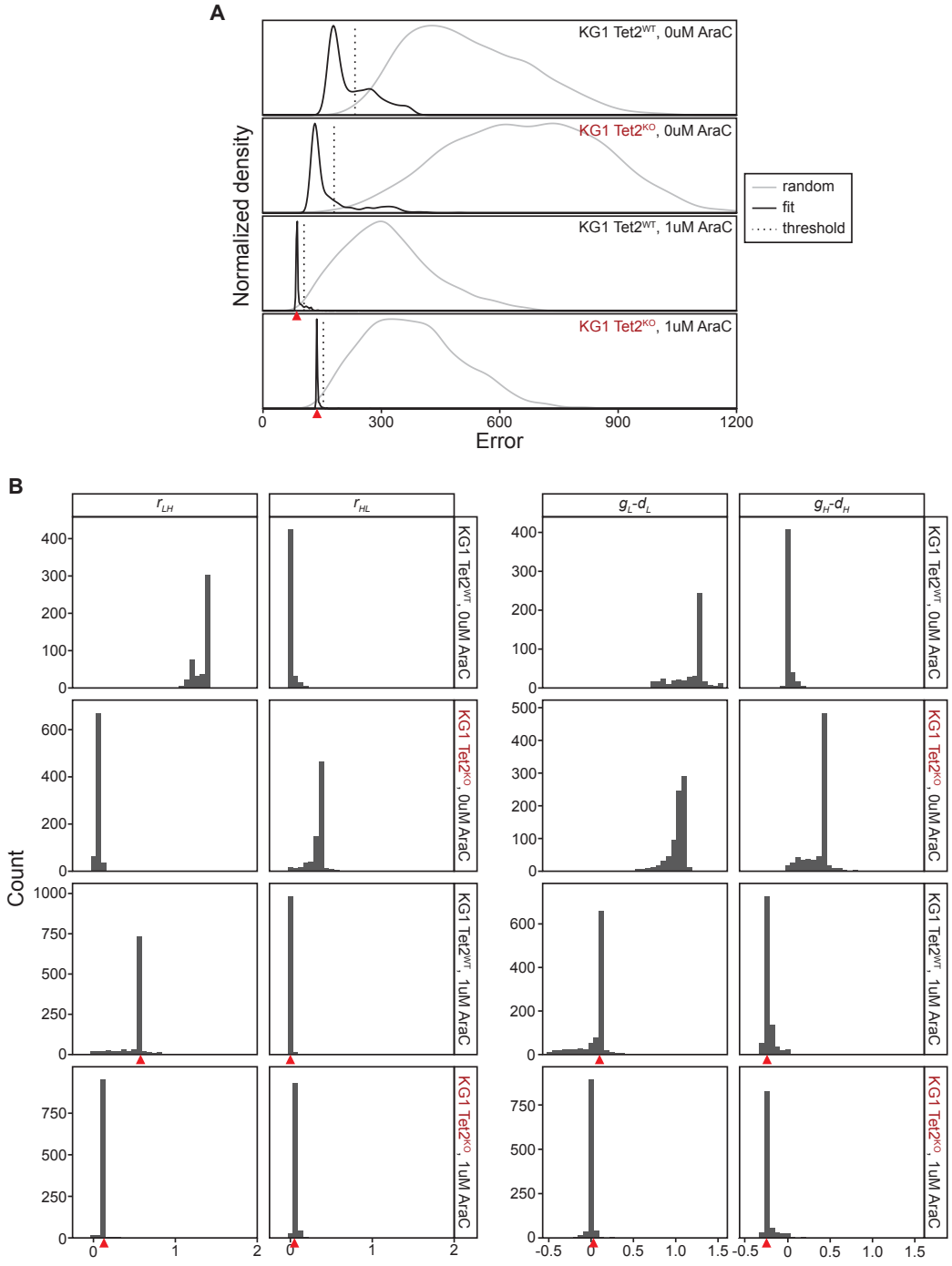

Figure A.10: Models converged to a small range of parameter values
(A) Model parameters were fit 1000 times for each condition with random initialization values. The distribution of prediction errors for fit model parameter values (black curve) are show as well as for random unfit parameter values (gray curve) for comparison. Solutions that had not converged were largely to the right of the threshold shown (vertical dotted line: 0.005-quantile for each random "null" distribution). (B) The distribution of parameter values for r_{LH}, r_{HL}, $net_L = g_L - d_L$ and $net_H = g_H - d_H$ are shown for solutions with error below the thresholds from (A). Most fit parameters have a single main value. Red triangles: error and parameter values shown in Figure 4.

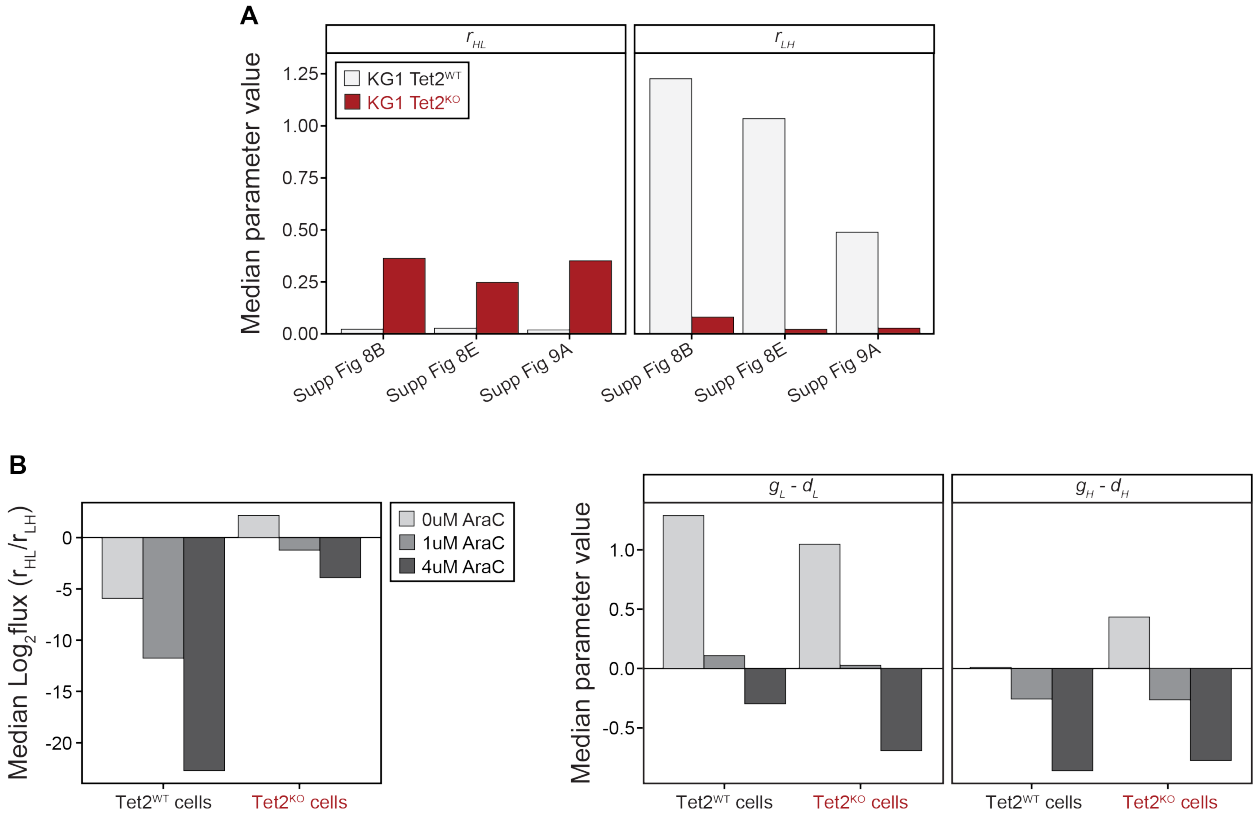

Figure A.11: Parameter values across experiments and drug conditions
(A) Shown are the median fit parameter values (see Methods) for the switching rates r_{HL} and r_{LH} out of drug for the 3 experiments shown in Figure A.8 (left), Figure A.8 (right), and Figure A.9. Even when sorting the "tails" (bottom/top 5%) of the CD38 distributions (as in Figure A.8), TET2^{KO} populations have a higher switching rate r_{HL} from H to L than their WT counterparts. (B) Shown are the median Log$_2$flux (where flux is defined as the ratio $^{r_{HL}}/_{r_{LH}}$) and median $g_L - d_L$ and $g_H - d_H$ values for cells treated with 0, 1, or 4μM AraC. While different conditions have different parameter values, Inequality 1 holds true ($g_L - d_L > g_H - d_H$) and TET2^{KO} populations have a higher H to L switching rate than TET2^{WT} ($Tet2^{KO}\frac{r_{HL}}{r_{LH}} > Tet2^{WT}\frac{r_{HL}}{r_{LH}}$).

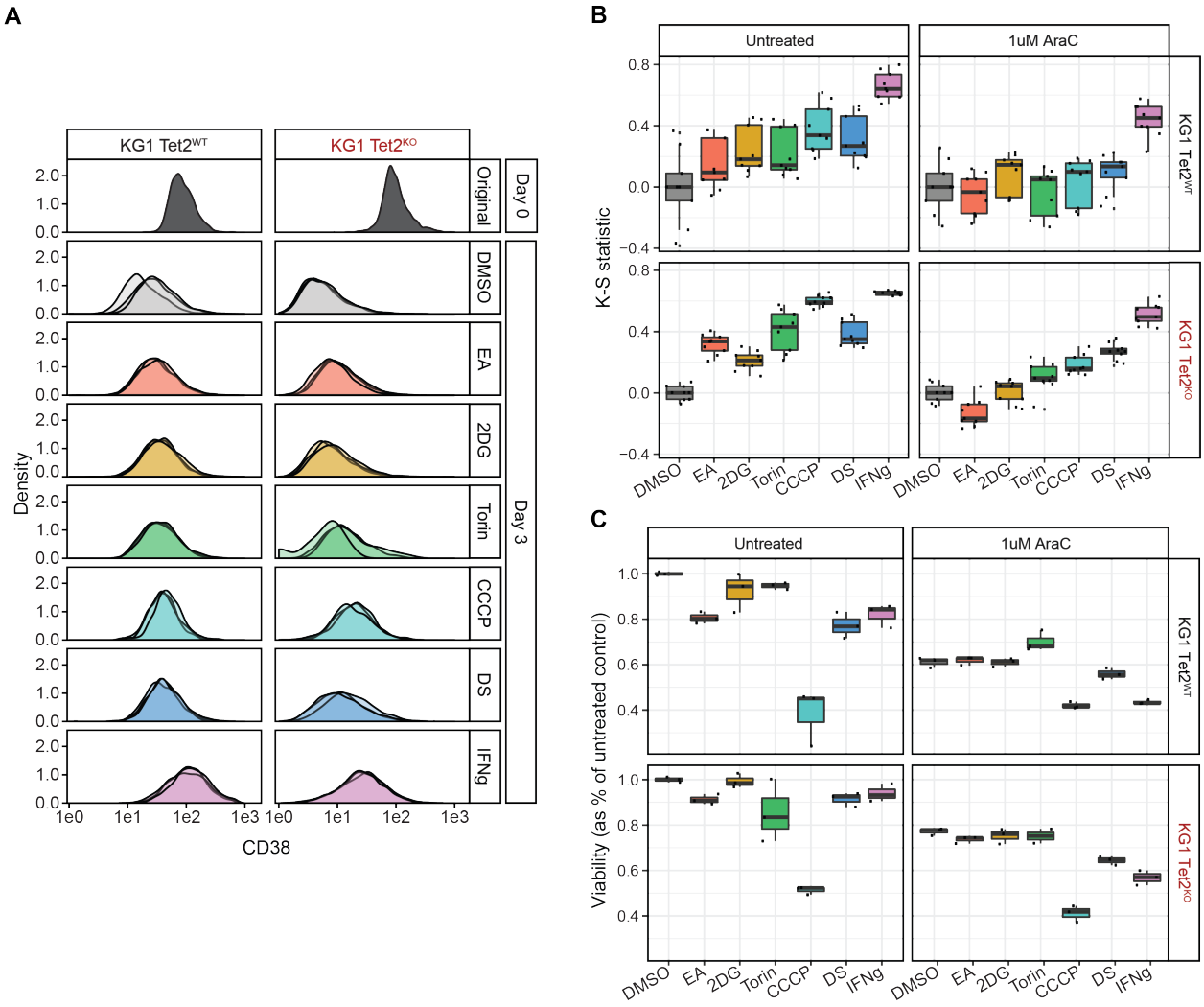

Figure A.12: Modulating state transitions and chemosensitivity with a panel of effectors
(A) KG1 and KG1 TET2^{KO} CD38^{hi} cells were isolated and plated in triplicate. The dark gray density shows the CD38^{hi} population reprofiled just after sorting. For 3 days, CD38^{hi} samples were treated with effectors, in the presence or absence of 1μM AraC treatment (untreated condition not shown). Samples were then reprofiled with flow cytometry to assess how CD38 surface marker expression had changed relative to DMSO control (light gray density, n=3). (B) The signed Kolmogorov-Smirnov statistic was calculated to quantify differences in CD38 surface marker expression in effector treatment relative to DMSO control (pairwise comparisons for n=3 technical replicates per condition). (C) KG1 and KG1 TET2^{KO} cells were treated with effectors, in the presence or absence of 1μM AraC treatment. After 3 days, viability was measured. Data shown is viability normalized to DMSO control in the untreated condition for each cell line (n=3 technical replicates per condition).

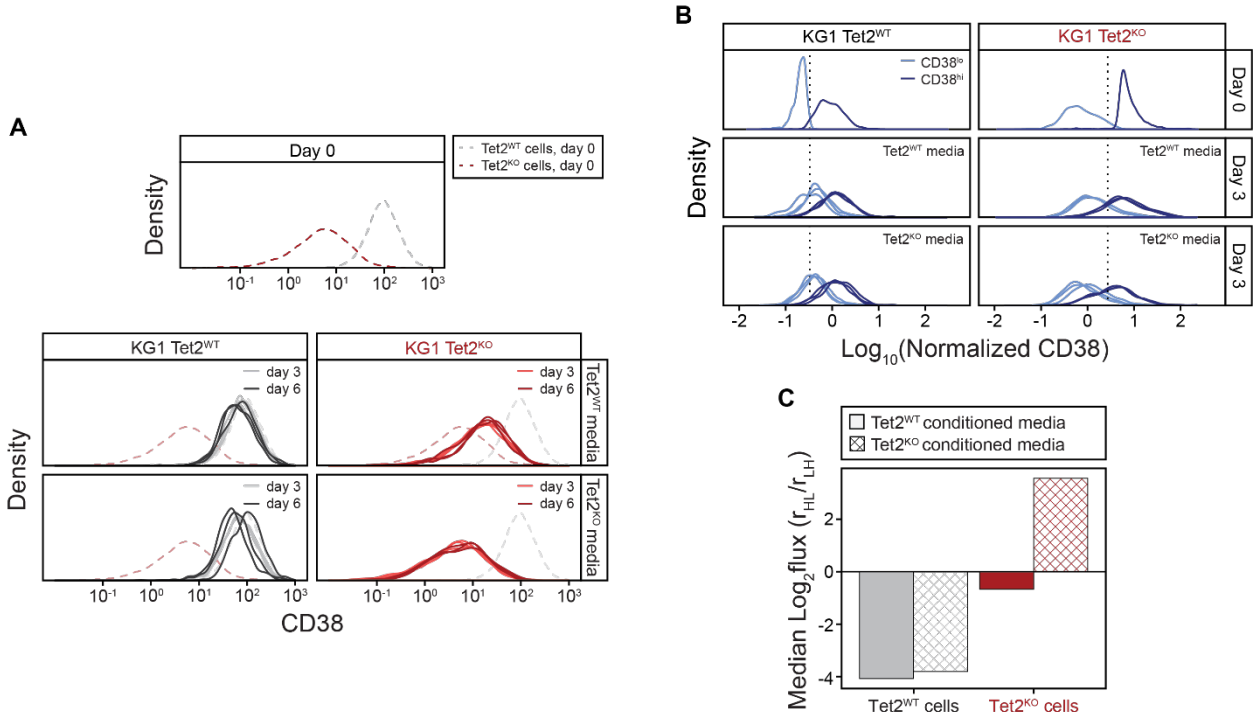

Figure A.13: Conditioned media from KG1 TET2^{WT} cells alters KG1 TET2^{KO} CD38 expression
(A) Unsorted TET2^{WT} and TET2^{KO} cells grown in either TET2^{WT} conditioned media or TET2^{KO} conditioned media were assessed for CD38 surface marker expression over the course of 6 days (n=3 per timepoint). While TET2^{WT} cells (left column) are not clearly affected by treatment with conditioned media from either cell line, TET2^{KO} cells (right column) have higher overall CD38 expression when grown in TET2^{WT} conditioned media. (B) TET2^{WT} and TET2^{KO} cells sorted into CD38^{lo} or CD38^{hi} (vertical dotted lines; see Figure A.5) were grown in either TET2^{WT} conditioned media or TET2^{KO} conditioned media and their CD38 surface marker expression was monitored for 3 days (n=3 per timepoint). (C) Shown are the median Log$_2$flux between the states L and H (where flux is defined as the ratio $^{r_{HL}}/_{r_{LH}}$) for TET2^{WT} and TET2^{KO} cells grown in either TET2^{WT} conditioned media or TET2^{KO} conditioned media (data from (B)). TET2^{WT} transition rates are not clearly affected by TET2^{KO} conditioned media, but TET2^{KO} transition rates show an increase in transition rate r_{LH} from L to H (lower $^{r_{HL}}/_{r_{LH}}$) in the presence of TET2^{WT} conditioned media.

Table A.1: Tissue expression profiles strongly correlated with TET2 expression in LAML study
Top Enrichr results from "ARCHS4 Tissues" for genes strongly correlated with TET2 expression (Pearson correlation > 0.45) in the LAML study.

Term	p-value	Adj p-value	Combined Score
GRANULOCYTE	5.80E-35	6.27E-33	351.221812
NEUTROPHIL	2.87E-26	1.55E-24	229.673308
DENDRITIC CELL	4.19E-23	1.51E-21	189.887693
PERIPHERAL BLOOD	1.37E-21	3.69E-20	171.757628
ALVEOLAR MACROPHAGE	2.09E-19	4.52E-18	146.67615
BLOOD DENDRITIC CELLS	6.88E-13	1.24E-11	80.1008484
MACROPHAGE	3.58E-11	5.53E-10	64.8236466
CORD BLOOD	1.28E-07	1.73E-06	36.6594651
SPLEEN (BULK TISSUE)	6.91E-06	8.29E-05	24.8371736
CD19+ B CELLS	9.91E-04	0.01070011	12.1748187
PLASMACYTOID DENDRITIC CELL	0.00389624	0.03825395	9.15439163

Table A.2: Tissue expression profiles strongly correlated with TET2 expression in OHSU study
Top Enrichr results from "ARCHS4 Tissues" for genes strongly correlated with TET2 expression (Pearson correlation > 0.45) in the OHSU study.

Term	p-value	Adj p-value	Combined Score
NEUTROPHIL	8.78E-22	9.49E-20	269.662309
GRANULOCYTE	1.18E-20	6.39E-19	248.485518
PERIPHERAL BLOOD	1.36E-13	4.90E-12	130.08582
DENDRITIC CELL	5.28E-11	1.42E-09	93.5226046
MACROPHAGE	5.84E-08	1.26E-06	56.0686849
ALVEOLAR MACROPHAGE	2.84E-07	5.12E-06	48.5367869
BLOOD DENDRITIC CELLS	1.29E-06	1.99E-05	41.6825623
CORD BLOOD	1.29E-06	1.74E-05	41.6825623
SPLEEN (BULK TISSUE)	0.00241362	0.02896349	13.2313735

Table A.3: Differentially regulated transcription factors from ChEA after $TET2^{KO}$
Top 15 Enrichr results from "ChEA 2016" for significantly differentially expressed genes in $TET2^{KO}$ AML cells compared to $TET2^{WT}$.

Term	p-value	Adj p-value	Z-score	Combined Score
RUNX1	2.33E-07	7.76E-06	-2.8484862	15.0816396
RARG	7.53E-05	0.00048047	-2.8505104	11.9960627
CEBPD	1.40E-09	3.82E-07	-1.5878123	9.32880721
RELA	1.81E-07	7.24E-06	-1.7199597	8.52545964
EBF1	1.91E-09	3.82E-07	-1.4942173	7.74789534
IRF8	0.00262177	0.00740246	-3.0702163	7.73525825
PPAR	1.31E-06	2.54E-05	-1.905316	7.71544316
PCGF2	0.00032077	0.00142563	-2.4373959	7.44619838
NFE2	9.37E-06	0.00010583	-2.0042818	7.2755795
JARID2	4.89E-06	6.24E-05	-1.9530783	7.05606928
SUZ12	2.82E-05	0.00023897	-2.1004067	6.99782768
CMYC	1.25E-08	8.31E-07	-1.5155208	6.70147917
FOXM1	1.25E-08	8.31E-07	-1.5138172	6.69394594
JUN	1.25E-08	8.31E-07	-1.5119283	6.68559372

Table A.4: Differentially regulated transcription factors from ENCODE after $TET2^{KO}$
Top 15 Enrichr results from "ENCODE and ChEA Consensus TFs from ChIP-X" for significantly differentially expressed genes in $TET2^{KO}$ AML cells compared to $TET2^{WT}$.

Term	p-value	Adj p-value	Z-score	Combined Score
RUNX1_CHEA	0.00159847	0.14546062	-1.6685561	10.1333255
GATA2_CHEA	0.01427494	0.43300653	-1.6725741	6.80445586
NFE2L2_CHEA	0.01198961	0.43300653	-1.6026109	6.72621148
IRF8_CHEA	0.02326501	0.47204626	-1.6126027	6.01136743
RELA_ENCODE	0.03356363	0.47204626	-1.6166295	5.30642849
TCF3_CHEA	0.02684883	0.47204626	-1.5232717	5.21913797
ESR1_CHEA	0.04234801	0.48170864	-1.5304939	4.77195633
FOXP2_ENCODE	0.05990861	0.54516831	-1.5211788	4.20867812
SUZ12_CHEA	0.03631125	0.47204626	-1.2969375	3.9741095
VDR_CHEA	0.05651921	0.54516831	-1.2799391	3.65467032
SMAD4_CHEA	0.07034123	0.57194911	-1.3844585	3.52724206
NANOG_CHEA	0.07542186	0.57194911	-1.3055066	3.2360662
TRIM28_CHEA	0.08775681	0.61429767	-1.2749751	3.03687525
NELFE_ENCODE	0.11161095	0.7254712	-1.3245068	2.83504397

Table A.5: Effectors used to halt the CD38^{hi} to CD38^{lo} transition

Name	Category	Max concentration
DMSO	vehicle	--
2-deoxy-D-glucose	glycolysis	2 μM
CCCP	oxphos	50 nM
Disulfiram	redox	1 nM
Etacrynic Acid	redox	100 nM
IFNg	signaling	50 ng/ml
Torin	AA	1 nM

Appendix B: Molecular effects of TET2 loss in AML cells

Results and Discussion

Growth advantage of TET2KO in KG1 cells

To understand the consequences of a TET2 mutation, we chose to compare isogenic human myeloblast cell lines expressing or not expressing a wildtype TET2. Isogenic KG1 cell lines were created by knocking out TET2 with Cas9-directed editing (see Chapter 2 Methods).

One week after nucleofection of the plasmid expressing Cas9 and our TET2-targeting guide (see Chapter 2 Methods), the region around the putative cut site was amplified via PCR and measured by Sanger sequencing. The Sanger traces were noisy beginning near the cut site (Figure A.1), implying a mixed population of edited mutants. The webtool TIDE was used to deconvolute the Sanger traces and confirmed a high likelihood of a mixed population including mutants and potentially unedited cells (denoted by bar at 0; Brinkman *et al*., 2014; Figure B.1, top). After three weeks of continuous passaging, the region near the cut site was again amplified and Sanger sequenced. TIDE deconvolution revealed that frameshift mutants (1bp deletion, 1bp insertion, or 2bp insertion; Figure B.1, bottom) that resulted in truncation upstream of catalytic domain (Figure A.1) comprised the remaining cell population, implying they outcompeted the rest of the cells seen on day 7. We did, in fact, notice an increased overall growth rate of KG1 $TET2^{KO}$ cells compared to $TET2^{WT}$, combined with an increased ability to continue growing at higher densities (Figure B.2). However, neither KG1 nor Thp1 showed a notable change in cell cycle latency (Figure B.3). While it is likely the growth rate measurement with CellTiter-Glo was capturing faster growth in addition to higher respiration in KG1 $TET2^{KO}$ cells (see Chapter 3), supplemental experiments counting cells with the TC20 automatic cell counter (Bio-Rad) and counting nuclei of Hoechst-stained cells confirmed a higher growth rate of KG1 $TET2^{KO}$ cells (data not shown).

Differential methylation patterns associated with TET2 loss

Previous work has shown that TET2 mutation leads to deficiencies in DNA demethylation (Yamazaki *et al*., 2015; Asmar *et al*., 2013). Targeted DNA methylation profiling was used to confirm that this held true in our isogenic cell lines. As mentioned in Chapter 2, $TET2^{KO}$ cell lines display a significantly higher degree of overall hypermethylation compared to their WT counterparts. Mean methylation across all

measured CpGs was not significantly altered in CpG islands in TET2^{KO} cells, consistent with literature (Rasmussen *et al*., 2015; Figure B.4, left). However, of the differentially methylated positions in KG1 TET2^{KO} cells, CpG islands were significantly hypermethylated compared to WT, and "shelf" and "open sea" CpGs were modestly, but significantly, hypomethylated (p-values 1e-7, 0.02, and 0.008 respectively; Figure B.4, right).

As mentioned in Chapter 2 and Appendix A (Figure A.4), we developed a score to assess epigenetic similarity to hematopoietic stem cells (HSCs) and LSCs (see Chapter 2 Methods) using a publicly available dataset of DNA methylation profiles from normal hematopoietic progenitors and leukemic cells (GEO GSE63409; Jung *et al*., 2015). I also wrote a similar method to assess transcriptomic similarity to HSCs using publicly available RNAseq data from sorted hematopoietic subpopulations (GEO GSE74246, Buenrostro *et al*., 2018). After normalizing the public dataset and our RNAseq data with ComBat in sva (version 3.10), I used the relative expression of genes and the distances between samples in transcriptome space as a proxy for sample similarity. Here, I show the expression of CD38 and ARHGEF17 by cell subpopulation and note that the TET2 mutants for both KG1 and Thp1 have more similar expression profiles to HSCs compared to their WT counterparts (Figure B.5, left). I also show the pairwise chord distance $d_{jk} = \sqrt{\sum_i \left(\sqrt{x_{ij}} - \sqrt{x_{ik}}\right)^2}$ of all samples to HSC samples, where the KG1 TET2^{KO} cells appear to be more similar to HSCs than TET2^{WT} (Figure B.5, right). Overall, these data further support the idea that TET2^{KO} populations acquire more stem-like signatures in KG1 and Thp1 cells.

TET2^{KO} may affect changes via transcription factor activity

In Chapter 2, we briefly discuss how analyses of both differentially methylated regions and gene set enrichment of the isogenic cell lines pointed to alteration of Runx1 (Methods), a master hematopoietic regulator known to affect fate decisions in the myeloid hierarchy. Specifically, the most significant differentially methylated region in Tet2^{KO} cells was the proximal promoter of Runx1 (minimum adjusted p-value 1.22e-27, Figure A.2), and Runx1 targets were highly enriched for differentially expressed genes in Tet2^{KO} cells (p-value 7.76e-6, Figure A.2, Tables A.3-A.4). Despite the Runx1 promoter being

significantly differentially methylated and expression of its target genes significantly changed, the RNA expression of Runx1 itself was not apparently different in either cell line (t-test, BH adjusted p-values 0.31 and 0.46). We therefore looked to Runx1 isoforms and whether their expression was altered in $Tet2^{KO}$ cell lines. Runx1 has proximal and distal promoter regions which initiate the expression of the isoforms Runx1a and Runx1c, respectively. Based on RT-qPCR, expression of Runx1a—but not Runx1c—in $Tet2^{KO}$ cells is strongly increased relative to the parental cell lines (fold-change 2.33, Figure A.2). Prior work has demonstrated that Runx1a is overexpressed in AML and other myeloid disorders (Liu *et al*., 2009; Sakurai *et al*., 2017), and that Runx1a expression expands hematopoietic stem cells (Tsuzuki *et al*., 2012; Ran *et al*., 2013). Thus, these data suggest that the increase in stem-like gene expression and colony-forming unit phenotypes in $TET2^{KO}$ cells is, in part, a result of differential methylation of a Runx1 promoter and differential expression of its isoform Runx1a. Future work studying the effect of Runx1 isoform overexpression on gene expression, drug survival, and self-renewal is necessary to test this claim.

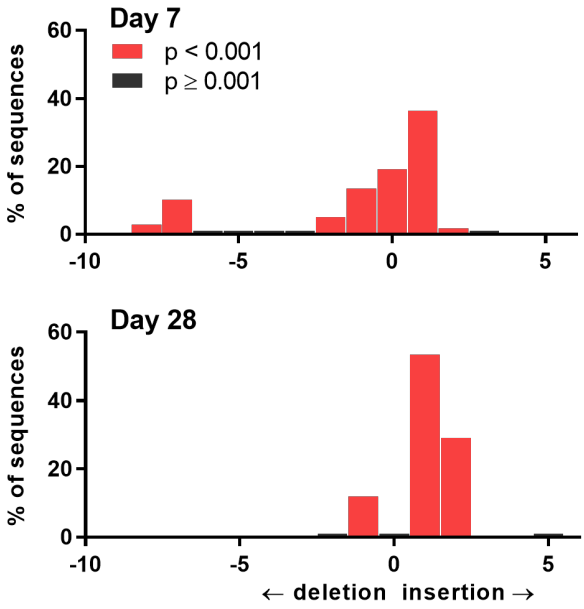

Figure B.1: Frameshift mutants dominate KG1 cells after TET2 editing
Deconvolution of Sanger sequencing traces with the webtool TIDE enabled estimation of mutant subpopulation prevalence within a bulk Cas9-edited KG1 population. Here, the x-axis represents subpopulations with deletions (-1bp, -2bp, etc), insertions (+1bp, +2bp, etc), or no change (0) detected in the Sanger sequencing trace. KG1 cells after Cas9-directed editing were grown for 1 week (top) and 4 weeks (bottom) prior to PCR amplification of the region near the edited locus and Sanger sequencing.

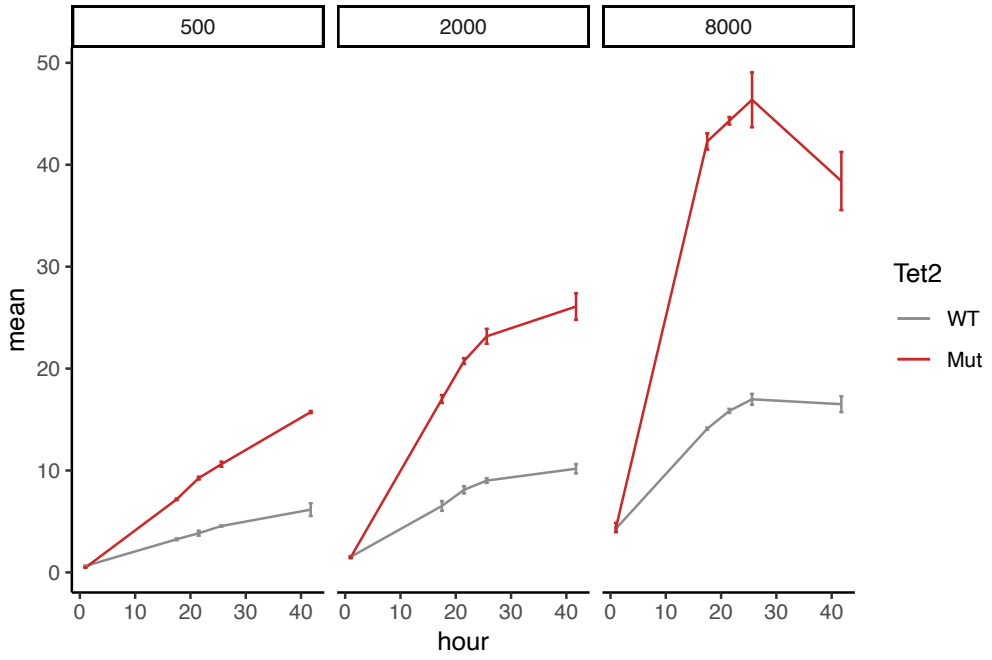

Figure B.2: KG1 TET2^{KO} cells grow faster than KG1 TET2^{WT}
KG1 cells were seeded at 500, 2000, or 8000 cells per well in 384-well plates, and cell viability was measured using RealTime-Glo over time. All luminescence values were normalized to the 500-cell wells at t=0. The x-axis represents time (hours) after seeding, the y-axis represents the normalized luminescence (AU; mean ± s.e., n=3).

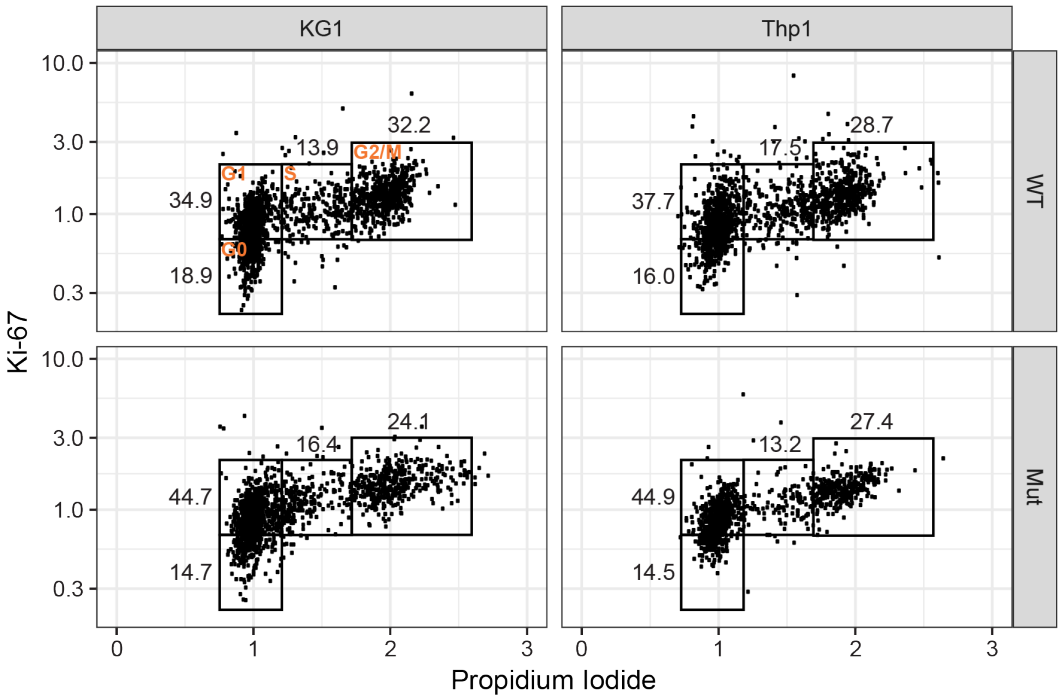

Figure B.3: Cell cycle latency is similar in TET2 mutants compared to WT
KG1 and Thp1 cells (both TET2 WT and mutant) were fixed and permeabilized, stained with Ki-67 antibody and propidium iodide, and assessed by flow cytometry for FITC and PI fluorescence (see Methods). Each point is a fixed cell, black boxes show the gates used for calling cells as G0, G1, S, or G2/M (orange labels), black text represents the percent of cells within the corresponding box.

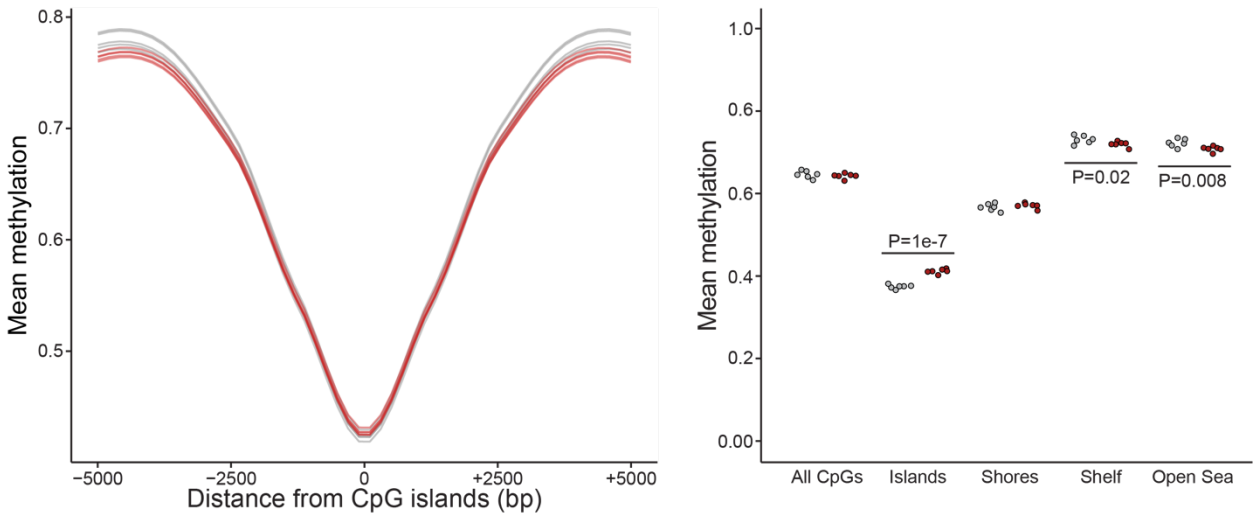

Figure B.4: Methylation changes of annotated CpGs in KG1 TET2^{KO} cells
(left) The beta value (methylation) of CpGs was plotted as a function of distance to its closest annotated CpG island (CpGi) start. Shown are loess smoothed means for KG1 TET2^{WT} (gray) and TET2^{KO} (red) cells for CpGs within 5000bp of a CpG island. (right) The mean beta value of differentially methylated CpGs by annotation shows that TET2^{KO} (red) samples have significantly higher methylation in CpGi compared to TET2^{WT} (gray).

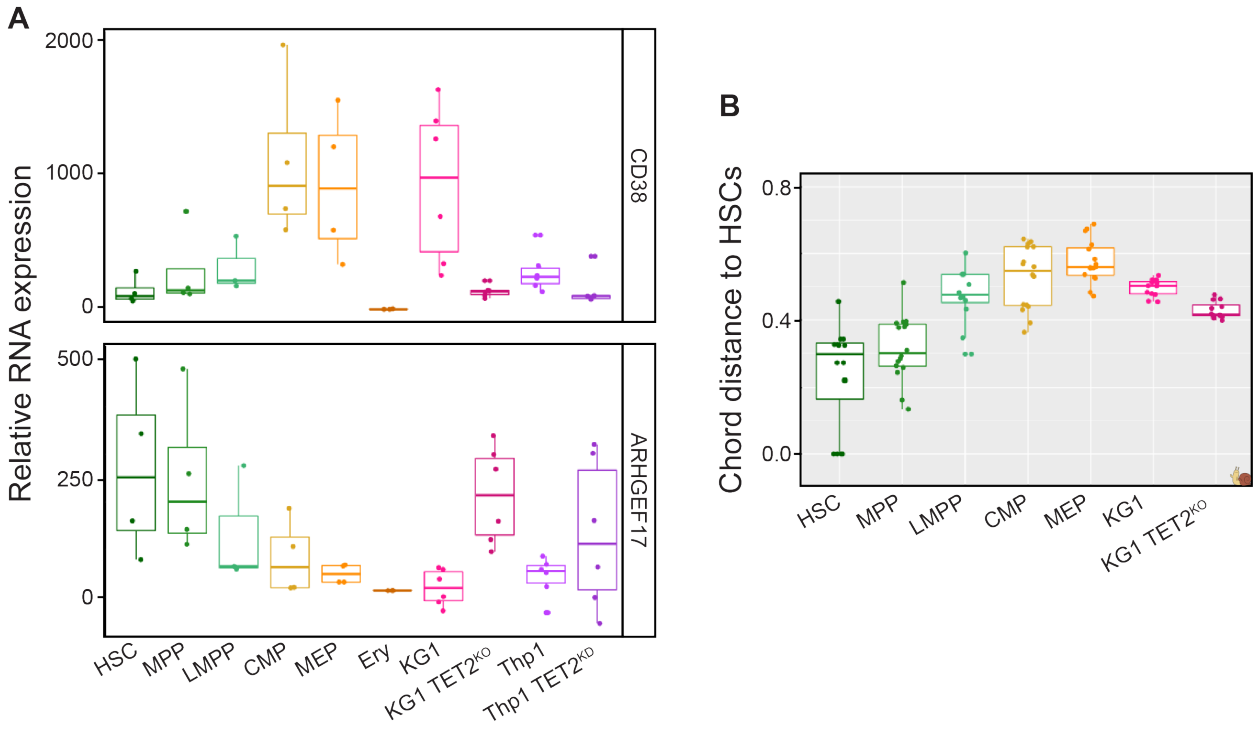

Figure B.5: TET2 mutants have more similar transcriptome profiles to HSCs than TET2^{WT}
(A) KG1 and Thp1 TET2 mutant expression of CD38 and ARHGEF17 is more similar to HSCs than their WT counterparts. Here RNA expression is shown for each sample compared to a publicly available dataset from sorted hematopoietic subpopulations (GEO GSE74246, Buenrostro *et al*., 2018). (B) The overall transcriptomic profiles of KG1 TET2^{KO} cells are closer to HSCs than KG1 TET2^{WT} (y-axis: pairwise chord distance to HSC samples).

Methods

Sanger sequencing and deconvolution

Isogenic KG1 cell lines were created by knocking out TET2 with Cas9-directed editing (see Chapter 2 Methods). The region surrounding the cut site was amplified via PCR (TET2 exon 3 F: 5'-TGTTTTTCCTGTGCCTGACCAG-3'; TET2 exon 3 R: 5'- GCTTCTGTGATTTGAGAGTAAGAGCC-3') with Jumpstart Taq ReadyMix (Sigma-Aldrich) according to manufacturer's protocol. Amplicons were purified using the DNA Clean & Concentrator-25 kit (Zymo Research) and sent for Sanger sequencing using the same forward primer.

Sanger traces were deconvoluted using the TIDE: Tracking of Indels by DEcomposition webtool (Brinkman *et al*., 2014; https://tide.deskgen.com) using default parameters.

Real-time cell viability assay

Cell viability was read out with RealTime-Glo MT Cell Viability Assay (Promega). Plates were prepared with MT Cell Viability Substrate and NanoLuc Enzyme as described in the manufacturer's protocol for "Continuous-Read Format: Reagent Addition at Cell Plating" in 384-well opaque white tissue culture plates (Corning). KG1 cells were then added at 500, 2000, or 8000 cells per well, and luminescence was read out with the Biotek H4 plate reader (BioTek, Winooski, VT) in the Center for Advanced Technology at UC San Francisco. Prior to each timepoint, plates were allowed to cool at RT for approximately 10 minutes. Luminescence values were normalized to the 500-cell wells at t=0 prior to visualization.

Assaying cell cycle status

KG1 and Thp1 cells were fixed and permeabilized according to a previously published protocol (Basic Protocol 1, Kim *et al*., 2015). Cells were stained with propidium iodide (Sigma-Alrich, #81845), Ki-67 (8D5) (Cell Signaling Technology, #9449), and Goat anti-Mouse Alexa Fluor 488 (Thermo Fisher Scientific, #A21131). Doublets were called based on gates drawn for FSC-A and FSC-W. All flow cytometry was performed on the Aria IIu in the Center for Advanced Technologies at UC San Francisco.

References

Asmar F, Punj V, Christensen J, Pedersen MT, Nielsen AB, Hother C, Ralfkiaer U, Brown P, Ralfkiaer E, Helin K, and Grønbæk K. (2013). Genome-wide profiling identifies a DNA methylation signature that associates with TET2 mutations in diffuse large B-cell lymphoma. Haematologica, 98(12), 1912–1920.

Buenrostro JD, Corces MR, Lareau CA, Wu B, Schep AN, Aryee MJ, Majeti R, Chang HY & Greenleaf WJ. (2018). Integrated Single-Cell Analysis Maps the Continuous Regulatory Landscape of Human Hematopoietic Differentiation. *Cell*, *173*(6), 1535-1548.e16.

Brinkman EK, Chen T, Amendola M & Van Steensel B. (2014). Easy quantitative assessment of genome editing by sequence trace decomposition. Nucleic Acids Research, 42(22), 1–8.

Liu X, Zhang Q, Zhang DE, Zhou C, Xing H, Tian Z, Rao Q, Wang M & Wang J. (2009). Overexpression of an isoform of AML1 in acute leukemia and its potential role in leukemogenesis. Leukemia, 23(4), 739–745.

Kim KH & Sederstrom JM. (2015). Assaying Cell Cycle Status Using Flow Cytometry. Current Protocols in Molecular Biology, 111(1), 139–148.

Ran D, Shia WJ, Lo MC, Fan JB, Knorr DA, Ferrell PI, Ye Z, Yan M, Cheng L, Kaufman DS & Zhang DE. (2013). RUNX1a enhances hematopoietic lineage commitment from human embryonic stem cells and inducible pluripotent stem cells. Blood, 121(15), 2882–2890.

Rasmussen KD, Jia G, Johansen JV, Pedersen MT, Rapin N, Bagger FO, Porse BT, Bernard OA, Christensen J, and Helin K. (2015). Loss of TET2 in hematopoietic cells leads to DNA hypermethylation of active enhancers and induction of leukemogenesis. Genes & Development, 29(9), 910–922.

Sakurai H, Harada Y, Ogata Y, Kagiyama Y, Shingai N, Doki N, Ohashi K, Kitamura T, Komatsu N & Harada H. (2017). Overexpression of RUNX1 short isoform has an important role in the development of myelodysplastic/myeloproliferative neoplasms. Blood Advances, 1(18), 1382–1386.

Tsuzuki S & Seto M. (2012). Expansion of functionally defined mouse hematopoietic stem and progenitor cells by a short isoform of RUNX1/AML1. Blood, 119(3), 727–735.

Yamazaki J, Jelinek J, Lu Y, Cesaroni M, Madzo J, Neumann F, He R, Taby R, Vasanthakumar A, Macrae T, *et al*. (2015). TET2 Mutations Affect Non-CpG Island DNA Methylation at Enhancers and Transcription Factor-Binding Sites in Chronic Myelomonocytic Leukemia. Cancer Research, 75(14), 2833–2843.

Appendix C

Appendix C: Supplemental Material for Chapter 3

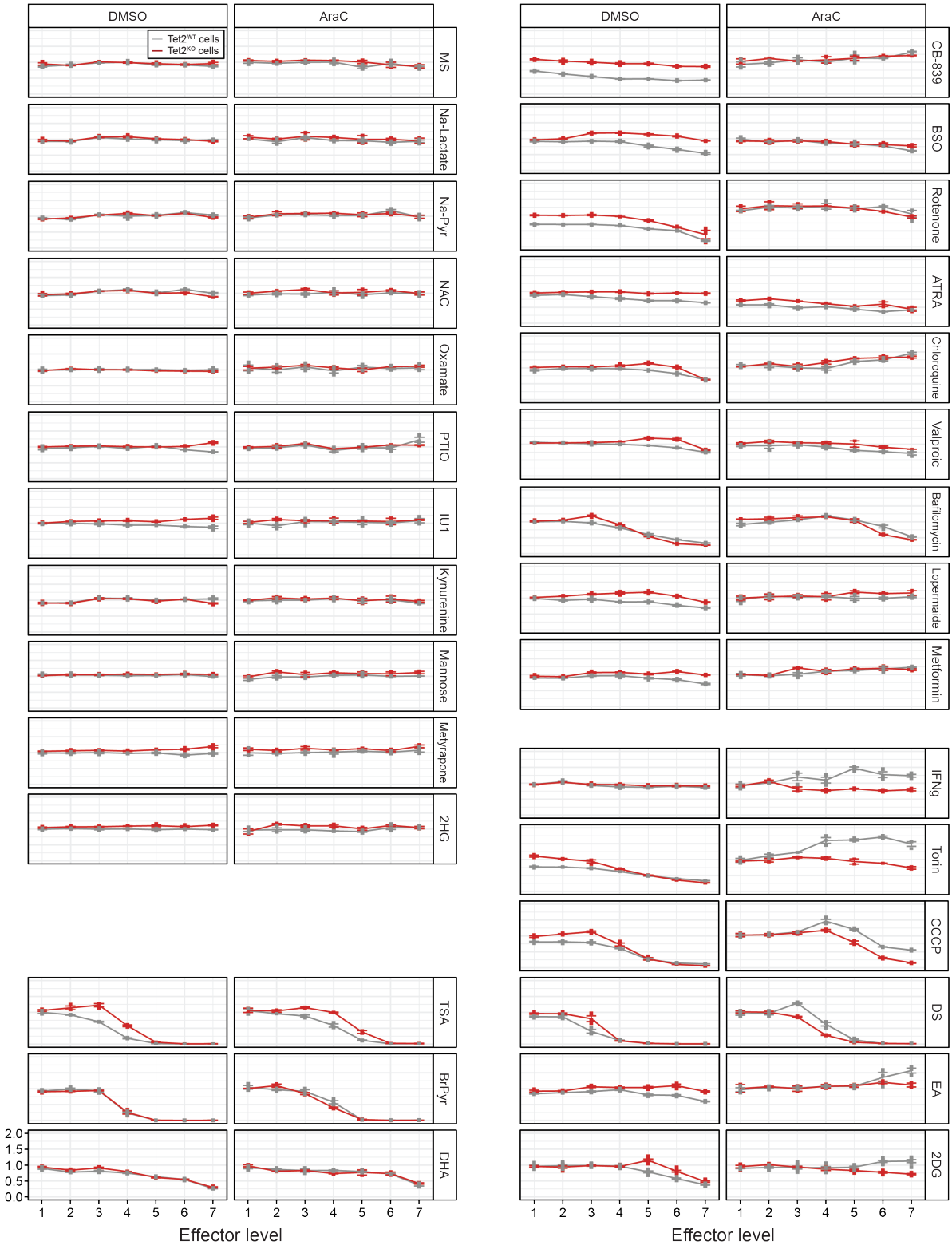

Figure C.1: Panel of effectors for modulating chemosensitivity
KG1 TET2^{WT} and TET2^{KO} cells were treated with a 2X dilution series of a panel of effectors (Table C.1) and either DMSO or 1μM AraC for 72h prior to cell viability measurement with CellTiter-Glo. Shown are dose-response curves (DCA, DNP, Fructose, MRT68921, and sulfasalazine not shown; mean ± s.e., n=3).

Table C.1: Panel of effectors used to modulate chemosensitivity

Name	Shorthand	Category	Max conc.
Dimethyl sulfoxide	DMSO	vehicle	--
2-deoxy-D-glucose	2DG	glycolysis	2 μM
2-hydroxyglutarate	2HG	oncometabolite	38.41 nM
All-trans retinoic acid	ATRA	retinoid	10 nM
Bafilomycin	--	autophagy	0.05 nM
Bromopyruvate	--	glycolysis	100 nM
Buthionine sulfoximine	BSO	redox	100 nM
CB-839		AA	10 nM
Carbonyl cyanide m-chlorophenyl hydrazone	CCCP	oxphos	50 nM
Chloroquine		autophagy	100 nM
Dehydroascorbate	DHA	redox	2 μM
Dichloroacetate	DCA	glycolysis	2 μM
Disulfiram	DS	redox	1 nM
Dinitrophenol	DNP	oxphos	200 nM
Etacrynic Acid	EA	redox	100 nM
Fructose		glycolysis	5.55 μM
IU1		autophagy	15 nM
Interferon gamma	IFNg	signaling	50 ng/mL
Kynurenine		AA	100 nM
Lopermaide		autophagy	10 nM
Mannose		glycolysis	5.55 uM
Mercaptosuccinate	MS	redox	100 nM
Metformin		oxphos	1 μM
Metyrapone		autophagy	100 nM
MRT68921		autophagy	1 nM
N-acetylcysteine	NAC	redox	5 μM
Na-Lactate		glycolysis	2 μM
Na-Pyruvate	Na-Pyr	glycolysis	2 μM
Oxamate		glycolysis	1 μM
PTIO		NO	100 nM
Rotenone		oxphos	1 nM
Sulfasalazine		redox/AA	100 nM
Torin		AA	1 nM
Trichostatin A	TSA	epigenetic	5 nM
Valproic Acid		epigenetic	200 nM

Publishing Agreement

It is the policy of the University to encourage open access and broad distribution of all theses, dissertations, and manuscripts. The Graduate Division will facilitate the distribution of UCSF theses, dissertations, and manuscripts to the UCSF Library for open access and distribution. UCSF will make such theses, dissertations, and manuscripts accessible to the public and will take reasonable steps to preserve these works in perpetuity.

I hereby grant the non-exclusive, perpetual right to The Regents of the University of California to reproduce, publicly display, distribute, preserve, and publish copies of my thesis, dissertation, or manuscript in any form or media, now existing or later derived, including access online for teaching, research, and public service purposes.

DocuSigned by:
Leanna Morinishi
591136D9D54045D...
Author Signature

5/22/2020
Date