Functional impact of inactivating mutations in epigenetic regulators in cancer

Ioannis Loukas

Imperial College London

And

The Francis Crick Institute

PhD Supervisor: Dr. Paola Scaffidi

A thesis submitted for the degree of

Doctor of Philosophy

Imperial College London, Department of Mathematics

March 2023

Statement of originality

I, Ioannis Loukas, hereby certify that the intellectual content of this thesis is the product of my own work and that all the assistance received in preparing this thesis has been acknowledged and relevant sources have been appropriately referenced.

Copyright declaration

*Θα ήθελα να αφιερώσω την παρούσα εργασία στον αδερφό μου **Κώστα** και στους γονείς μου **Ασημίνα** και **Δημήτρη** για την απεριοριστη στηριξη τους*

Abstract

Cancer evolution is driven by selection acting on genetic and epigenetic diversity to promote the propagation of the fittest subpopulations. This phenomenon is shaped by the tumor microenvironment which is often characterized by stressful conditions. Epigenetic regulators are frequently mutated during the later stages of tumorigenesis, but the functional impact of their inactivation is poorly understood.

In this thesis, I hypothesize that the disruption of the epigenetic regulatory network increases cell fitness in unfavorable environments and thus is selected over time. Through large-scale fitness assays in various cancer models, I demonstrate that epigenetic deregulation leads to a widespread stress-specific survival advantage. This effect is mediated by mutations in all layers of epigenetic regulation, is shared across different stress conditions and is cancer type independent. Then, I explore various cellular mechanisms that can underlie this stress-specific fitness advantage. Genetic diversity, transcriptional heterogeneity or phenotypic plasticity cannot explain the increased survival under stress, as revealed by a combination of reversible epigenetic inhibition, live-cell imaging and single-cell transcriptomics. On the contrary, epigenetically deregulated cells remain phenotypically inert (less responsive) under stress. Transcriptional profiling of cancer populations in hostile conditions, revealed significant alterations in fitness and growth-related signatures. Disruption of the epigenetic machinery results in a defective stress response, thus decreasing the probability of such cells to surpass a stressed threshold and ultimately die. This defective transcriptional rewiring underpins the inert phenotype that emerges upon epigenetic deregulation.

Collectively, by investigating the effect of inactivating mutations in epigenetic regulators on cell fitness under environmental stress, I propose that phenotypic inertia is the favorable cellular trait that is selected over time. My findings provide a potential explanation for the widespread subclonal mutations affecting epigenetic regulators and have significant implications for cancer evolution.

Acknowledgements

I would like to thank my PhD supervisor Dr. Paola Scaffidi, who gave me the opportunity to work within her lab for the past five years on this demanding but ultimately exciting project. Her critical thinking regarding biological phenomena, guidance over the years and the constant challenging of my thoughts significantly shaped my scientific thinking and where I stand today as a researcher.

Being part of the Cancer Epigenetics Lab has been a great experience. I was honored to be surrounded by totally different, yet kind and brilliant people who all together fostered an outstanding environment. Cristina, Josep, Tom and Elanor welcomed me into the lab and since then they have been invaluable inside and outside of it. I have really enjoyed being around them. Special thanks must go to Cristina, who was always there for us providing all her knowledge, expertise and literally anything we asked. She is a stellar example of how a scientist should be and the lab would not be the same without her! A big thank you must go to Paolo, that was involved in this study. I feel lucky that I got the chance to interact with him, learn from him and work together towards significant discoveries. Puay has been my lab alter ego. Being side-by-side in the lab in the first couple of years, we shared good experiences, bad experiences, and literally countless hours working together. It has been a real journey getting to know her. A special thank you must also go to Thomas (The Robustness Guy). Brilliant but humble, he was always there to spark spirited conversations on the floor of 3NW quadrant, constantly challenging my thoughts and making questions that most of the times I did not have the answers to. Some of them were related to science too! Finally, I want to thank Fabrizio and Marta that followed up on this work and contributed immensely towards our discovery. Many thanks also to the rest of the current and past Scaffidi lab members!

Credit must also go to my thesis committee members along with all the people from the Crick STPs and the 3NW quadrant along with fellow researchers from Crick and Imperial College London that I interacted with over the years. Their help, input and advice were invaluable.

Science aside, I was lucky enough to get to know some people that I will call friends for life. People that we shared too many memories to even count and that I consider them my family from London! Nikos, Michele, Foteini, Spyridoula, Tatiana, Elina are few of them to mention. Along with my family they had always been there for me during the good but most importantly the rocky parts of this adventure. Finally, huge thanks should go to Margarita for her support during the writing of this thesis. I would not have done it without her.

To all the people named above and the ones not mentioned by mistake,

A huge huge thank you!

Table of Contents

Table of Figures

List of Tables

Abbreviations

AML	Acute Myeloid Leukemia
Amp	Ampicillin
ASNS	Asparagine Synthetase
ATAC	Assay for Transposase-Accessible Chromatin
BET	Bromodomain and Extra-Terminal
BSD	Blasticidin
Capture sequence	Specific oligonucleotide that was cloned in the 3' end of the sgRNA scaffold, enabling the capture of expressed sgRNAs from the 10X Genomics single cell transcriptomics pipeline
Cas9	CRISPR-Associated Protein 9
CDK9	Cyclin Dependent Kinase 9
cDNA	Complementary DNA
ChIP	Chromatin immunoprecipitation
Chl	Chloramphenicol
cMYC	MYC Proto-Oncogene
Cntr	KO cells for TNP2, a non-expressed gene
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats
CTC	Circulating Tumor Cell
CTD	Carboxy-Terminal Domain (of RNA polymerase II)
CV	Coefficient of Variation
DDR	DNA Damage Response
DEGs	Differentially Expressed Genes
Dox	Doxycycline
EED	Embryonic Ectoderm Development
EGFP	Enhanced Green Fluorescent Protein
EMT	Epithelial to Mesenchymal Transition
ERs	Epigenetic Regulators
EZH2	Enhancer Of Zeste 2 Polycomb Repressive Complex 2 Subunit
FACS	Fluorescence-Activated Cell Sorting
FDR	False Discovery Rate
FRET	Förster Resonance Energy Transfer
gDNA	Genomic DNA
GG	Golden Gate
GLUL	Glutamine Synthetase
GOF	Gain of Function

GSEA	Gene Set Enrichment Analysis
H&E	Hematoxylin and eosin
H_20_2	Hydrogen peroxide
HCl	Hydrochloric Acid
HDAC	Histone Deacetylase
HFGs	High-Frequency Genes
HU	Hydroxyurea
HVGs	Highly-Variable Genes
IDH	Isocitrate Dehydrogenase
IF	Immunofluorescence
IHC	Immunohistochemistry
IL1B	Interleukin 1 Beta
Indels	Insertions/Deletions
IP	Intraperitoneal
KDM	Lysine Demethylase
KMT	Lysine Methyltransferase
KO	Knock-out
KP model	Cancer models driven by oncogenic mutations in *TP53* and *KRAS*
LB	Luria Broth
LOF	Loss Of Function
LQ cells	Low Quality cells; as defined in the scRNA-seq analysis
LUAD	Lung adenocarcinoma
LUSC	Lung Squamous Cell Carcinoma
ML	Maximum Likelihood
mRNA	Messenger RNA
NGS	Next Generation Sequencing
NLS	Nuclear Localisation Signal
NSG mice	NOD.Cg-Prkdc^{scid}Il2rg^{tm1Wjl}/SzJ Immunocompromised mice
OXPHOS	Oxidative Phosphorylation
PCA	Principal Component Analysis
PCR	Polymerase Chain Reaction
PDX	Patient-Derived Xenograft
PDX L1C5c	Derivative clone from LXFL 1674, inducibly expressing Cas9
PDX LXFL 1674	Patient derived xenograft cell line; Large Cell Carcinoma
PDX MeA5a	Derivative clone from MEXF 2090, inducibly expressing Cas9
PDX MEXF 2090	Patient derived xenograft cell line; cancer type: Melanoma

PFA	Paraformaldehyde
PRC2	Polycomb Repressive Complex 2
qPCR	Quantitative PCR
RPS26	Ribosomal Protein S26
RT	Reverse Transcribed
rxn	Reaction
SCLC	Small Cell Lung Cancer
SD	Standard Deviation
SEM	Standard Error of the Mean
sgRNA	Single guide RNA
SLC1A3	Solute Carrier Family 1 Member 3
SMARCD1	SWI/SNF Related, Matrix Associated, Actin Dependent Regulator Of Chromatin, Subfamily D, Member 1
SNVs	Single Nucleotide Variation
SUZ12	SUZ12 Polycomb Repressive Complex 2 Subunit
TIDE	Tracking of Indels by DEcomposition
TME	Tumor Microenvironment
TMRE	Tetramethylrhodamine, Ethyl Ester, Perchlorate
TNP2	Transition Protein 2
TPM	Transcripts Per Million
TRACERx	TRAcking Cancer Evolution through therapy (Rx)
UMAP	Uniform Manifold Approximation and Projection
UMI	Unique Molecular Identifier
VAF	Variant Allele Frequency
WES	Whole Exome Sequencing
WT	Wild Type

Chapter 1. Introduction

1.1 Conceptualization of cancer as an evolutionary disease

More than a century ago, zoologist and anatomist Theodor Boveri studied mechanisms of inheritance by observing chromosome segregation during mitosis in sea urchins. In one of his latest works, he made several predictions about tumorigenesis, suggesting that cancer cells arise from normal cells and that alterations in chromosome number is the root cause of the disease (Boveri, 2008). He also proposed that such abnormalities can lead to the loss of growth-inhibitory chromosomes and/ or the amplification of growth-promoting chromosomes, ultimately driving the uncontrolled proliferation phenotype observed in canner (Boveri, 2008). Pioneering studies in the following decades confirmed most of his predictions regarding tumorigenesis, focusing on its genetic basis (Wunderlich, 2002).

In the late 1970s, Peter Nowell proposed the clonal nature of cancer and established it as an evolutionary disease. He suggested that tumorigenesis occurs in a linear manner, where genetic diversity within a single clone drives the selection and domination of subsequent sublines (Nowell, 1976). Nowell also proposed various models that could explain the emergence of such genetic alterations, distinguishing among cell-intrinsic events like the aberrant activation of genomic loci, and cell-extrinsic factors like environmental mutagens (Nowell, 1976). For many years, tumorigenesis was considered a linear process that progressed in a stepwise manner through the gradual acquisition within a single dominating clone of favorable mutations.

In the past two decades, technological advances in next-generation sequencing and single-cell profiling methodologies allowed for a thorough investigation of the mutational landscape at different stages during tumor progression. These studies revealed substantial clonal heterogeneity within all cancer types (McGranahan & Swanton, 2017; Nam, Chaligne & Landau, 2021). Subsequent in silico inference of their evolutionary dynamics demonstrated that population heterogeneity can be the result of distinct processes like linear, branched or even neutral evolution (Vendramin, Litchfield & Swanton, 2021). In their seminal study, Gerlinger et al. performed exome sequencing in multiple regions within the same tumor and provided initial evidence of branched evolution, with spatially separated clones carrying distinct favorable driver mutations that are selected over time (Gerlinger et al., 2012; Vendramin, Litchfield & Swanton, 2021). On the other hand, other studies proposed that the detected clonality is a result of early events that drive an explosive increase in intratumor heterogeneity, followed by relaxed selective pressures allowing for the propagation of such

populations over time (neutral evolution) (Sottoriva et al., 2015). The classification of cancers into the above distinct categories can be influenced by the experimental methodology applied and the time point used to infer the evolutionary history of a tumor. Thus, the most probable scenario is that all modes of evolution co-exist at distinct stages during tumorigenesis (Turajlic et al., 2019).

In summary, over the past century the view of how cancer evolves has significantly shifted. The initial notion that cancer progresses linearly has now been replaced by the realization that cancer is a collection of highly heterogeneous diseases that can be characterized by distinct evolutionary dynamics even within the same patient (Turajlic et al., 2019). Despite this variability, the conceptualization of cancer as an evolutionary disease allows its description through a framework that examines the key forces that drive evolutionary phenomena in asexual populations (Lipinski et al., 2016) [Figure 1], namely:

I. *Population diversification*: stemming from genetic and epigenetic alterations along with their interaction with the tumor microenvironment.
II. *Selection*: acting on population heterogeneity to promote the propagation of the fittest.
III. *Drift*: describing alterations in gene variant frequency due to random chance.

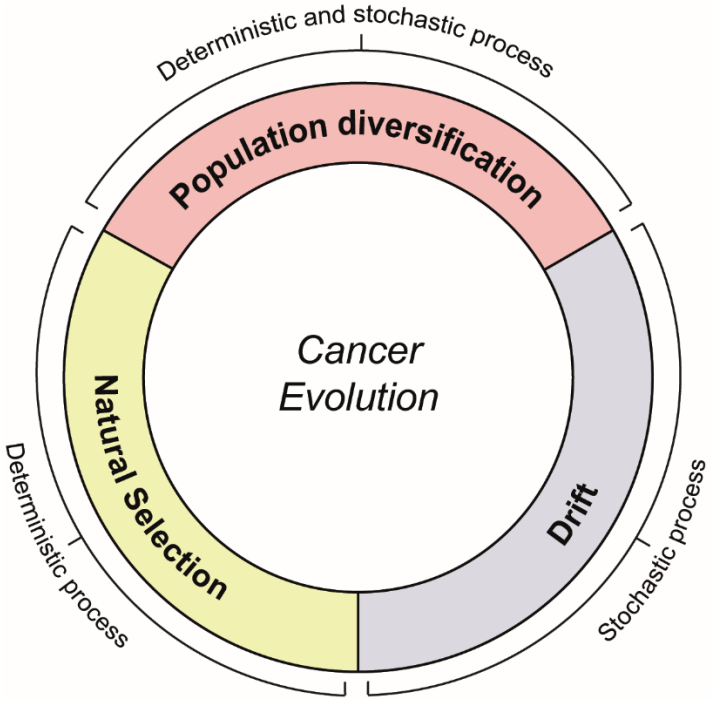


Figure 1: Studying cancer through an evolutionary lens

Modified from Lipinski et al (2016)

The following subchapters of the Introduction will explore the first two forces in greater detail, starting with population diversity shaped by genetic alterations.

1.2 Genetic alterations as a driver of cancer evolution

1.2.1 Types, causes and consequences

The best characterized source of diversity within a cancer population is alterations at the genomic level. Sequencing studies have revealed a wide spectrum of abnormalities that operate at different scales, affecting single genes and their regulatory elements or larger portions of the genome. These include single-nucleotide variations (SNVs), small insertions/deletions (indels), copy-number alterations (CNAs) and numerical or structural chromosomal alterations (Burrell et al., 2013; McGranahan & Swanton, 2017). The extent of the genetic diversity and the nature of the detected mutations across cancer types vary significantly, reflecting distinct cell-intrinsic and cell-extrinsic mechanisms at play (Dentro et al., 2021). For instance, pediatric and hematological malignancies that often emerge from single oncogenic alterations are characterized by a small mutational burden. Conversely, melanoma and lung cancer reside on the other side of the spectrum reflecting their development after years of exposure to environmental mutagens like UV irradiation and tobacco respectively (Burrell et al., 2013; McGranahan & Swanton, 2017; Dentro et al., 2021). Aside from such environmental factors, a plethora of cell-intrinsic defects in pathways involved in DNA replication, DNA damage repair and chromosomal segregation have been proposed as sources of intra- and inter-tumor genetic variability (Burrell et al., 2013).

While increased genetic diversity is associated with worse clinical prognosis (Morris et al., 2016), not all mutations in a tumor are functional. Based on their impact on the survival and proliferation capacity of cancer cells, these alterations can be classified as *passenger* (no effect) or *driver* (positive effect). Recent advances in mathematical modelling (Dentro et al., 2021), multi-region sequencing (Gerlinger et al., 2012) and single-cell genomics (Saadatpour et al., 2015) provided insights into the temporal order of their acquisition, identifying mutations that are either *clonal* or *subclonal* based on their presence in all or a fraction of the cancer cells respectively. For example, mutations that inactivate various tumor suppressor genes like *TP53* and *CDKN2A* or activate oncogenic drivers like *KRAS* are frequent clonal events across multiple cancer types (Dentro et al., 2021). Upon these strong initiating hits that promote malignant transformation, cells acquire further genetic diversity that promotes the acquisition of diverse oncogenic related features (Hanahan, 2022) that fuel disease progression and relapse upon therapy (McGranahan & Swanton, 2017; Hanahan, 2022). Considering that these late-occurring mutations were only recently identified, their functional impact on subclonal expansion is largely understudied.

1.2.2 Genetic abnormalities cannot explain all aspects of cancer evolution

Despite the incrimination of hard-wired genetic alterations as drivers of cancer evolution, several experimental observations cannot be explained by such events. Genome sequencing of normal tissues, like skin and oesophageal epithelium, demonstrated that normal cells can carry somatic mutations at comparable levels to cancerous lesions. Additionally, they detected significant positive selection and clonal expansion of cells carrying oncogenic mutations in various cancer driver genes, including *TP53, NOTCH1* and *KMT2D* (Martincorena et al., 2015, 2017, 2018). The fact that such mutations did not lead to the acquisition of malignant potential suggests that either additional genetic hits are required to initiate tumorigenesis or that non-genetic / microenvironmental factors play a crucial role in determining the likelihood of this event.

Genomic analysis across various cancer types revealed a lack of driver clonal mutations in 20% of the examined cases and only a minor fraction of subclones carrying driver events in the form of SNVs or indels (Dentro et al., 2021). In pancreatic cancer, comparison of the mutational landscape between matched primary and metastatic samples, showed a remarkable similarity in the genetic composition, failing to reveal novel drivers of metastasis (Yachida et al., 2010; Hu et al., 2020). Similar results were also obtained by a recent pan-cancer study of metastatic tumors (Priestley et al., 2019). Although these observations may reflect the suboptimal detection of low-prevalence genetic drivers within the primary tumor, an alternative scenario is that additional non-genetic factors govern these behaviors.

In the clinical setting, targeted inhibition of EGFR or BRAF in lung cancer and melanoma patients respectively, leads to rapid initial remission followed by relapse. Exome sequencing of MAPK-treated melanoma samples upon relapse, did not detect any known genetic resistance drivers in 40% of such cases (Hugo et al., 2015). On top of that, it has been reported that cancer cells can regain their sensitivity to targeted therapy after drug withdrawal for a selected period of time (Sun et al., 2014). Such reversible phenotypic behaviors are inconsistent with the resistance being genetically encoded. Adding on the notion that mutations cannot always explain the respective phenotype of cancer cells, the temporal order of acquisition of driver mutations in *JAK2* and *TET2* in patients suffering from Myeloproliferative neoplasms was shown to drastically influence the evolutionary trajectory of the disease. Depending on the order in which the mutations were acquired, cells presented differences in multiple properties like their capacity to initiate tumorigenesis or resist to therapy (Ortmann et al., 2015).

Collectively, abnormalities within the genome fall short of explaining various observations regarding tumor initiation, maintenance and relapse, suggesting that additional layers of regulation are pivotal in shaping the phenotypic behaviors of cells during disease progression. Such non-genetic determinants of cancer evolution are explored in the following section.

1.3 The emerging importance of non-genetic mechanisms in cancer

1.3.1 Brief introduction to epigenetic regulation

From a historical perspective, the field of epigenetics emerged in order to explain a complex set of cellular behaviors that generate phenotypic diversity without underlying changes in the genotype (Allis & Jenuwein, 2016). The most notable manifestation of epigenetics is found in organismal development, where the same DNA sequence gives rise to multiple distinct cell types with diverse morphologies and functions. To collectively describe these molecular processes that explain the decoupling between genotype and phenotype, in 1942 the embryologist Conrad Waddington introduced epigenetics as "the branch of biology which studies the causal interactions between genes and their products which bring the phenotype into being." (Waddington, 2012). Since then, multiple variations of the original definition of epigenetics emerged, drifting away from its strict association with development, in order to describe experimental observations within distinct fields of research that are not necessarily overlapping conceptually (Deans & Maggert, 2015). In 1994, Holliday suggested that epigenetics reflects "the study of the changes in gene expression, which occur in organisms with differentiated cells, and the mitotic inheritance of given patterns of gene expression." (Holliday, 1994). This definition introduced an important aspect of epigenetic regulation, which is associated with the establishment of stable gene expression states and cellular inheritance. In the early 2000s an updated definition emerged that incorporated elements from all the distinct school of thoughts mentioned above. Based on this consensus definition, epigenetics reflects the "the study of changes in gene function that are mitotically and/or meiotically heritable and that do not entail a change in DNA sequence" (Wu Ct & Morris, 2001; Berger et al., 2009).

The above definition of epigenetics (or epigenetic regulation) is broad as it portrays multiple different regulatory layers that can lead to variability in the observed phenotypes that is not genetically encoded. Non-genetic or epigenetic mechanisms responsible for such diversification include, but are not limited to, signalling, alterations within the chromatin, long non-coding RNAs, covalent modification of the transcriptome, splicing, translational regulation

and many more. Here, I focus on the core panel of epigenetic regulators that directly act on chromatin.

Chromatin is the substrate upon which epigenetic activities introduce functional and structural variation. These activities act on multiple distinct layers, involving DNA methylation, a variety of histone modifications, alterations in nucleosome occupancy, histone variants and regulation of higher-order conformation [Figure 2]. These properties are not randomly distributed across the genome but rather are located in specific locations (e.g. proximal or distal regulatory elements, gene body, etc). In a location-dependent manner, the complex interaction of chromatin modifications, epigenetic regulators and downstream effectors orchestrates the transcriptional output of genes along with packaging of the chromatin fiber itself (Shen & Laird, 2013; Allis & Jenuwein, 2016).

The core principle of epigenetic regulation is based on the presence of 3 classes of proteins, termed *writers*, *erasers* and *readers* (Allis & Jenuwein, 2016). Writers carry enzymatic activities for the deposition of covalent modifications on DNA and histones, while readers recognize these modifications and recruit further effector proteins to “translate” or propagate the epigenetic state. These modalities make epigenetic information a heritable trait that can be maintained over cell duplication (Almouzni & Cedar, 2016). Erasers, on the other hand, mediate the removal of these modifications, underpinning the second important aspect of epigenetic regulation, its reversibility. The stability of chromatin states is of particular interest in cancer as it allows for the selection and propagation of favorable traits that are encoded epigenetically. Through their reversible nature, and in contrast to hard-wired mutations, epigenetic mechanisms can mediate alterations in the phenotype of cancer cells in response to various internal or external stimuli.

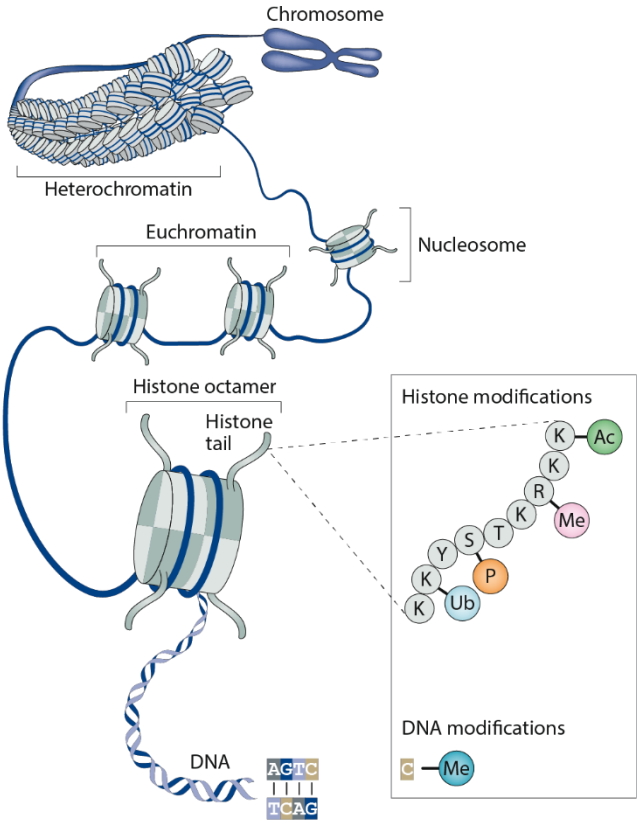


Figure 2: Layers of core epigenetic regulation

Epigenetic mechanisms include various enzymatic and accessory activities that act on the chromatin. Epigenetic proteins belong to distinct functional classes including DNA and histone modifiers, histone variants, chromatin remodellers and regulators of higher-order organization. By altering functional and structural properties within chromatin, epigenetic regulators orchestrate the transcriptional output from specific loci and from a developmental perspective are implicated in establishing cellular identity.

Me: Methylation, Ac: Acetylation, P: Phosphorylation, Ub: Ubiquitylation, K: lysine
Adapted from Hogg et al (2020)

In the literature, experimental observations that do not have an apparent genetic basis have been described as non-genetic or epigenetic, often in an interchangeable manner. Here, the term *non-genetic* describes cases of phenotypic deviation with no detectable genetic basis, such as transcriptional variance in genetically identical cells. On the other hand, *epigenetic* is used in a more restricted manner to describe experimental evidence that directly explore specific properties within the chromatin fiber, like histone modifications and nucleosome occupancy. In the following subchapter, both non-genetic and epigenetic determinants of cancer evolution are discussed with an emphasis on the latter.

1.3.2 Non-genetic and epigenetic abnormalities in cancer cells

In addition to mutations, a plethora of epigenetic abnormalities have been described to contribute to the heterogeneity observed in cancer cells. The best characterized epigenetic alteration in cancer is DNA methylation, a stable modification often associated with gene repression. Cancer cells routinely exhibit widespread loss of intergenic DNA methylation along with a focal increase of methylation in proximal regulatory elements (Baylin & Jones, 2011). Notably, intra-tumor variability within the DNA methylation landscape is predictive of patient outcome, suggesting its importance for cancer (Mazor et al., 2016). During tumorigenesis, the post-translational modifications of histones get substantially deregulated and are also prognostic of patient outcome (Kurdistani, 2007). Compared to DNA methylation, their pattern within tumors is less understood. However, emerging data revealed significant variability within cancer populations that is important both for tumor progression (McDonald et al., 2017) and relapse to therapy (Marsolier et al., 2022). Structural components of the chromatin itself, have been also observed to be differentially expressed across cancer cells, with the most notable case being H1.0. Torres et al. reported substantial intra-tumor variability in the expression levels of this linker histone and demonstrated that its differential regulation leads to alterations in various properties, like the self-renewal capacity of cancer cells (Torres et al., 2016). Finally, aside from the core epigenetic marks and proteins, recent sequencing studies at the single-cell level have uncovered substantial transcriptional heterogeneity even within cells with an identical genetic background (Nam, Chaligne & Landau, 2021).

The emergence of epigenetic and transcriptomic variability within cancer cells can be attributed to various cell-intrinsic and extrinsic sources (Li, Seehawer & Polyak, 2022). First, cancer cells have the inherent ability to transit between different states (Brock, Krause & Ingber, 2015). For instance, in their pioneering study, Gupta and colleagues separated a breast cancer line into its constituents (basal, luminal and stem-like states) and demonstrated that each state can give rise through phenotypic switching to a heterogeneous population resembling the parental one (Gupta et al., 2011). Second, some cancer types have been shown to follow a hierarchical organization, where a fraction of cells demonstrate stem-cell like properties and fuel disease progression whereas the rest of the population consists of non-tumorigenic differentiated cells (Wainwright & Scaffidi, 2017). As the chromatin state is linked with the differentiation status, it is expected that developmental hierarchies in various cancers (AML, *IDH*-mutant glioblastoma, etc) can contribute to the observed diversification of the epigenetic and transcriptomic patterns. Third, the chromatin landscape can be influenced by other regulatory layers within the cells like genomic and metabolic alterations (Shen & Laird, 2013; Flavahan, Gaskell & Bernstein, 2017). Finally, tumor cells do not grow in isolation but rather in a highly dynamic and heterogeneous environment (TME) that can significantly shape

the phenotypic properties within cancer cells (TME is discussed in greater detail in section 1.4.2).

1.3.3 Non-genetic and epigenetic mechanisms as drivers of cancer evolution

1.3.3.1 Tumor initiation and progression

It is now well appreciated that non-genetic and epigenetic alterations can significantly contribute to the tumorigenic potential of cells. Here I discuss few representative paradigms that act at different scales.

One characteristic case is the epigenetic silencing of genes. As discussed earlier, imbalance within the DNA methylome is a frequent phenomenon in cancer, often manifesting by hypermethylation in CpG islands proximal to the transcription start site of genes (Baylin & Jones, 2011). As DNA methylation in promoter regions leads to repression, these focal enrichment events result in the epigenetic silencing of genes, in an analogy to gene inactivation by mutations. Indeed, these events are frequent in cancer and have been reported to affect multiple tumor suppressor genes including regulators of the cell cycle (e.g. *RB1*, *CDKN2A*), lineage-specific transcription factors (*GATA4/5* in colon cancer), DNA Damage response proteins (e.g. *BRCA1, FANCF*) and mediators of apoptosis (*TRAILR1*) (Baylin & Jones, 2011; Shen & Laird, 2013). Although some dispute regarding the causality of these events exists, epigenetic silencing of genes via DNA hypermethylation or other epigenetic activities is considered a widespread phenomenon that fuels both tumor initiation and progression (Kazanets et al., 2016).

On the other hand, focal hypermethylation can indirectly lead to aberrant gene expression during tumorigenesis. An example of this is THE methylation of CTCF binding sites, leading to loss of insulation along the 3D genome and promiscuous interactions within regulatory elements (Chaligne et al., 2021). For example, in *IDH*-mutant gliomas characterized by hypermethylation, Flavahan and colleagues observed an abnormal contact between a constitutive enhancer and the regulatory elements of the *PDGFRA* oncogene. Reversing the epigenetic state was sufficient to rescue CTCF binding and repress the oncogenic driver (Flavahan et al., 2016). The above phenomenon is representative of cases where local alterations in strict epigenetic control may lead to the aberrant expression of singles genes that drive tumorigenesis.

From a systems perspective, stemming from organismal development, epigenetics can be considered as the safeguard of cellular identity, a property that is considerably challenged during tumorigenesis. In non-cancerous contexts, epigenetic mechanisms act as a barrier to the loss of cellular identity and reprogramming (Kim et al., 2021). Thus, in an analogy between

cellular reprogramming and tumorigenesis, the epigenetic landscape within a cell can act as a barrier to malignancy by affecting the susceptibility to oncogenic stimuli. As a proof of principle, in a zebrafish model of tumorigenesis, ectopic expression of BRAFV600E oncogene in differentiated melanocytes and their respective progenitors resulted in vastly different phenotypes, with the malignant transformation being significantly more probable when the oncogene was expressed in stem cells (Baggiolini et al., 2021). Another mechanism via which epigenetic abnormalities fuel tumor initiation is by directly hindering the faithful execution of proliferation-differentiation decisions in cells. A plethora of mutations in diverse epigenetic regulators in both solid and haematological malignancies have been shown to disrupt differentiation and stemness related programs, thus providing evidence regarding the causal relationship between alterations in epigenetic regulatory layers and the acquisition of tumorigenic capacity. Such genetic abnormalities affecting epigenetic proteins are described in greater detail in chapter 1.5.1.

Focusing back on tumor progression, variability within cancer populations is one of the most well appreciated properties of cancer cells. Many terms have been used to describe experimental observations of diversity like plasticity, transcriptional heterogeneity, lineage infidelity, state transitions and many more. Although there are fundamental conceptual differences in the above terms (see below), their commonality resides in the description of phenotypic diversity within a cancer population that enables the adoption of various malignant-promoting properties (Hanahan, 2022). The most well studied case of plasticity in cancer is the epithelial-to-mesenchymal transition (EMT). EMT is a developmental program that involves the acquisition of mesenchymal characteristics from originally epithelial cells. This process, along with its intermediate states, is co-opted by cancer cells to drive various aspects of cancer evolution (Nieto et al., 2016).

Significant insights regarding the importance of non-genetic and epigenetic variability in cancer progression have emerged from recent studies that take advantage of the technological advances in single-cell profiling and lineage tracing (Nam, Chaligne & Landau, 2021). These studies have provided at least three major advances in the field of cancer evolution. First, the simultaneous exploration of DNA, epigenetic and transcriptomic properties within individual cells revealed many instances where the detected variability does not have a genetic basis, thus suggesting that such variability is an inherent property of cancer cells worth further investigation (Fennell et al., 2022). On top of that, similar approaches provided a first line of evidence supporting the notion that transcriptomic variability has an epigenetic basis, implicating DNA methylation, chromatin accessibility and histone modifications as potential drivers (Corces et al., 2016; Pastore et al., 2019; Johnson et al., 2021). However, the correlative nature of these studies fails to draw causal links as discussed in greater detail in

section 1.5.3. Second, transcriptional variance within the population can be heritable. In their pioneering study, Fennel et al. provided evidence of stable non-genetic mechanisms as drivers of cancer progression. One of their observations involved the detection of transcriptional signatures that could be mitotically inherited and were driving clonal dominance (Fennell et al., 2022). A complementary approach detected similar stability of rare transcriptional states within cancer populations. Interestingly, these variable signatures were enriched among others for response to stress and signalling suggestive of the importance of such fluctuations for tumor progression (Shaffer et al., 2020). Third, they provided a methodological advance for distinguishing between plastic and transcriptionally variable subpopulations, two distinct scenarios that have been largely explored interchangeably so far (Mills, Stanger & Sander, 2019). The confusion mainly arises from the inability of static data to infer the temporal evolution within the system and assess if the observed variability stems from state transitions or selection over time from a heterogeneous pool of transcriptionally stable cells. Overlaying the single-cell profiling data on the lineage relationships of cells, inferred from the DNA methylation history (Chaligne et al., 2021) or evolvable barcodes (Yang et al., 2022), can circumvent this problem. By using such an approach, Chaligne et al. demonstrated that two distinct categories of brain cancers (*IDH*-wt and *IDH*-mutant gliomas) are characterized by significantly different state transition dynamics during disease progression. These studies provide a framework for future exploration of how cell state transitions contribute to various aspects of tumor evolution.

The above discussion is by no means exhaustive, but rather a selected presentation of examples at distinct scales that demonstrate the importance of non-genetic and epigenetic mechanisms in defining the malignant properties pivotal for cancer initiation and progression. Metastasis and resistance to therapy are two major evolutionary bottlenecks during carcinogenesis and thus are explored separately in the following sections.

1.3.3.2 Metastasis

One major event during disease progression is metastatic dissemination. During this multi-step process, cells detach from the primary tumor, circulate through the bloodstream, colonize new tissues and subsequently after a period of dormancy establish secondary tumors (Fares et al., 2020). Lineage tracing of cells during metastasis coupled with single-cell transcriptomics revealed cases where cells belonging to the same clone demonstrated drastically distinct metastatic potential (Quinn et al., 2021). On top of that, the variable expression of metastasis associated gene signatures among these clonal populations suggests the contribution of non-genetic mechanisms to the process (Nguyen et al., 2016; Quinn et al., 2021). Similarly, the

emergence of a rare cancer subpopulation marked by the expression of SEMA3C in genetically identical cells has been recently proposed to drive metastasis in a melanoma model of tumorigenesis (Kaur et al., 2022).

Various lines of correlative and functional studies have provided evidence that epigenetic alterations in DNA methylation, histone modifications and chromatin accessibility can act as key regulators of metastasis (Chen & Yan, 2021). Exome sequencing of matched primary and untreated tumors in breast, colorectal and lung cancer patients revealed that the driver mutations restricted to metastasis were significantly enriched for enzymes involved in chromatin binding and modification, indicating an important role of the epigenetic regulatory layer towards metastatic dissemination (Hu et al., 2020). In the case of pancreatic cancer, similar studies showed similar genetic landscapes between primary and metastatic lesions. (Makohon-Moore et al., 2017). Conversely, comparison of the epigenetic landscape during PDAC progression revealed significant alterations in both histone marks (methylation and acetylation) and DNA methylation during metastasis (McDonald et al., 2017). Specifically, distant metastatic lesions in the lung or liver were characterized by the loss of methylation marks both in histones (H3K9me2 and H3K27me3) and the DNA, that facilitated the upregulation of signatures affecting diverse cellular traits like metabolism and locomotion. Interestingly, a small fraction of cells within the primary tumor presented similar chromatin alterations to the metastatic one, indicative of a case where a rare pre-existing epigenetic state was selected over time during metastasis (McDonald et al., 2017). In pancreatic cancer, alterations in the enhancer activity driven by the pioneer transcription factor FOXA1, allowing for the hijacking of pro-metastatic developmental programs, have been proposed as an alternative non-genetic route towards metastasis (Roe et al., 2017). ATAC-sequencing of primary and metastatic samples in small-cell lung cancer (SCLC) patients revealed dramatic differences in the accessibility of intergenic regions enriched for the binding of NFIB. This transcription factor was heterogeneously expressed in the primary tumor and was sufficient to drive the “opening” of the chromatin that facilitated the upregulation of genes involved in various aspects of the metastatic cascade (Denny et al., 2016). Collectively, alterations within the chromatin landscape that allow for the acquisition of various transcriptional programs have been proposed as drivers of metastasis.

1.3.3.3 Relapse to therapy

Building on the observation that cancer cells can naturally transit between distinct cell states (Gupta et al., 2011), Sharna and colleagues were the first to demonstrate the presence of a slow-cycling cell state in NSCLC cells that survives upon EGFR inhibition and mediates

therapy resistance (Sharma et al., 2010). On the other hand, melanoma cells exposed to concurrent RAF/MEK inhibition exhibited a gradual transition between distinct transcriptional states. This process involved the early emergence of a stress-induced “starved” phenotype, followed by the adoption of diverse trajectories towards resistance that were associated with the activation of either differentiation or stemness related programs (Rambow et al., 2018). The above cases represent two distinct evolutionary frameworks where resistance is conferred either by the Darwinian selection of pre-existing resistant states (Sharma et al., 2010) or by the Lamarckian induction of favorable ones (Rambow et al., 2018). Multiple studies have demonstrated that these scenarios are not mutually exclusive, as evident by the detection of rare primed cells within cancer populations that are initially selected upon exposure to stress and subsequently undergo genetic or epigenetic reprogramming to achieve stable resistance to therapy (Shaffer et al., 2017; Hong et al., 2019). Currently, a plethora of non-genetic mechanisms has been proposed as mediators of therapy resistance, including transition towards a slow proliferation state, alterations in lineage identity or metabolic rewiring (Boumahdi & de Sauvage, 2020; Shen, Vagner & Robert, 2020).

Alterations in chromatin properties and factors have been associated with the emergence of drug-tolerant cancer cells. Upon treatment of diverse cancer models with therapy, cells exhibit common alterations in various chromatin marks. In detail, the drug-tolerant state is often characterized by loss of H3K4me3 and H3K27me3 (active and repressive chromatin marks respectively) and a concordant gain in H3K9me3 that resides within heterochromatic regions (Sharma et al., 2010; Ravindran Menon et al., 2015; Liau et al., 2017; Emran et al., 2018). These alterations are accompanied by differential expression of the respective writers and erasers, like for example lysine methyltransferases (KMTs; EZH2, SETDB1, etc) or demethylases (KDM; KDM5B, KDM6A, etc). The emergence of the slow-cycling tolerant state is also associated with reduced DNA methylation levels, driven by the upregulation of TET2 (Puig et al., 2018). In their recent study, Marsolier and colleagues detected a fraction of treatment-naïve cancer cells that exhibit a similar epigenetic state, driven by alterations in the H3K27me3 landscape, to the favorable one upon treatment, providing an example where a pre-existing epigenetic state is selected and enriched upon therapeutic intervention (Marsolier et al., 2022). Finally, genetic ablation or pharmacological inhibition of distinct epigenetic regulators with diverse catalytic activities is sufficient to tune the emergence of the tolerant state underscoring the importance of epigenetic mechanisms in the process (Sharma et al., 2010; Ravindran Menon et al., 2015; Liau et al., 2017; Emran et al., 2018; Puig et al., 2018; Marsolier et al., 2022).

1.4 Selective pressures during cancer evolution

1.4.1 Fitness is a dynamic cellular property

In the previous chapters, I have explored how cell-intrinsic (genetic and non-genetic) alterations contribute to population diversity, upon which selection acts to promote the survival and the propagation of the fittest subpopulations. The term *cellular fitness* can be used to describe the ability of a cell to thrive in a given environment (Di Gregorio, Bowling & Rodriguez, 2016). In quantitative terms, this can be described by the number of surviving offspring in a given time frame. The above definition establishes fitness as an intrinsic cell property that is dependent on i) the proliferation capacity of a cell per se and ii) the ability of this cell and its progeny to survive external challenges.

Considering its dependency on environmental factors, fitness is not fixed across the evolutionary trajectory of the disease but is rather dynamic and dependent on external fluctuations. Thus, it is expected that a phenotypic trait, acquired through a genetic or epigenetic alteration, can be favorable in a specific timeframe during cancer progression but dispensable later. Similar fluidity in cell fitness can be expected also in the spatial dimension. During tumor progression, such alterations in the fitness landscape are more prominent during significant evolutionary bottlenecks where there is a drastic alteration in the selective forces such as during metastatic dissemination or therapeutic intervention (Ciriello & Magnani, 2021). In line with that, in their recent study, Salehi et al. utilized genomic analysis of triple-negative breast cancers at the single-cell level followed by mathematical modelling to infer fitness from clone frequencies at different time points and conditions (Salehi et al., 2021). In response to chemotherapy, the thriving subclones were distinct from the ones that were dominating the fitness landscape under unperturbed conditions. On top of that, upon withdrawal of the stressor, the resistant clones collapsed indicating a fitness cost of resistance to therapy (Salehi et al., 2021). Similar shifts in the fitness landscape were also observed in melanoma, where multi-colour lineage tracing revealed the presence of distinct cancer subpopulations that drive primary growth and metastatic dissemination (Karras et al., 2022). Accordingly, in a mouse model of lung cancer, the metastatic potential was shown to be uncorrelated with the clonal size and the proliferative capacity of the cells (Quinn et al., 2021). Aside these major evolutionary events, it is reasonable to speculate that milder selective forces within the tumor microenvironment can shape what is considered favorable and thus the selective phenomena observed within cancer patients.

1.4.2 The components of the tumor microenvironment (TME)

To gain a deeper understanding of the selective forces that shape the evolution within the tumor core, it is essential to examine the components that make up the tumor microenvironment (TME) [Figure 3]. The TME can be broadly characterized by its cellular components along with its physical and chemical properties. The former category describes all the non-malignant cells that can be found within a tumor core and includes among others stromal, endothelial and infiltrating immune cells. Cancer-associated fibroblasts and immune cells present extensive intra- and inter-tumor heterogeneity and a growing body of evidence suggests that they can exert both pro- and anti-tumorigenic effects during disease initiation and progression (Binnewies et al., 2018; Sahai et al., 2020). Such functional interactions between cancer and non-malignant cells are not the scope of this thesis and are not explored further. In physiological conditions, cells rely on a well-organized network of vessels to receive sufficient amounts of nutrients, oxygen and pro-survival molecules. However, during the establishment of solid tumors, the uncontrolled proliferation of cells leads to significant disruption of tissue homeostasis (Almagro et al., 2022). The tumor vasculature specifically is characterised by poor architecture and functionality, leading to a suboptimal supply of key resources. The resulting acellular hallmarks of the tumor microenvironment are hypoxia, nutrient scarcity and acidosis (increased extracellular pH as a result of hypoxia and the glycolytic activity of cancer cells) (Wei et al., 2020).

Analysis of hypoxia across various cancer types has revealed significant variability both between patients and across cancer types. For instance, squamous cell tumors of the head and neck or lung were the most hypoxic, whereas thyroid and prostate adenocarcinomas resided on the other side of the spectrum (Bhandari et al., 2019). Even within a single tumor, the physicochemical properties are not homogeneous. Indeed, the disorganized and spatially heterogeneous vasculature and the metabolic activity of cancer cells can significantly shape the regional patterns of oxygen, nutrient availability and acidity (Wei et al., 2020). The most notable manifestation of such regional diversity can be found when comparing the outer and inner fractions within cancer lesions. The tumor core is often characterized by poor vascularization, nutrient scarcity and increased levels of apoptotic or growth-arrested cells. Conversely, the tumor periphery is associated with increased oxygen levels, acidic conditions and a proliferative phenotype of the cells (Rohani et al., 2019). Thus, regional heterogeneity in resources alters the local selectable forces at play and thus may shape subclonal expansions observed during tumorigenesis. For example, in a resource-rich region, subclones with higher intrinsic proliferative capacity will expand. By contrast, in locations characterized by hostile conditions, subclones that are able to survive and adapt to these stressors will outcompete their counterparts and become the prevalent population. How the

physicochemical properties within the TME change over time is less studied, but it is expected that the fitness landscape (what is considered favorable) will also vary in this dimension.

To sum up, cancer cells do not grow in isolation but rather in complex communities with non-malignant cells, in an environment characterized by hostile conditions. These properties are spatially and temporally heterogeneous posing constant novel challenges to tumor cells. This variability within the TME creates diversity in the local selective forces thus shaping what is considered a favorable and selectable trait during evolution. Considering the importance of TME in tumorigenesis, the effects of genetic or epigenetic perturbations on cellular fitness should be explored in various contexts reminiscent of the ecological niche of cancer cells (Zahir et al., 2020).

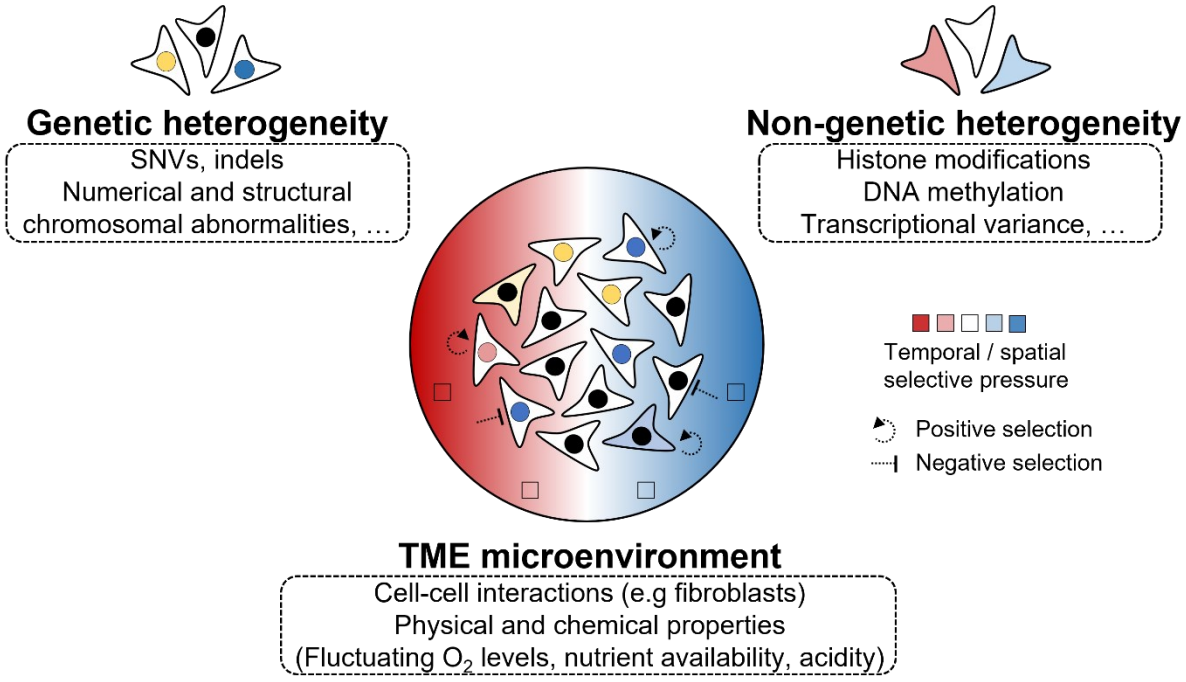


Figure 3: Fitness is a dynamic property shaped by the tumor microenvironment

The tumor bulk is composed by a heterogeneous population of cells harboring distinct genetic and epigenetic profiles (depicted here with diverse colors). Darwinian selection favors the survival of the fittest clones (arrows), a notion that is shaped by the tumor microenvironment, which is spatially and temporally dynamic.
Pre-existing or *de novo* acquired genetic/epigenetic traits fuel disease progression by enabling cells to survive and expand within their hostile setting.

1.5 Widespread mutations affecting diverse epigenetic regulators

1.5.1 Overview of the mutational landscape

Epigenetic regulation operates on multiple layers and involves the stable yet reversible modification of various properties within chromatin (section 1.3.1). While this strict regulation is important for cellular function, it gets significantly challenged during carcinogenesis with functional consequences for disease progression. This deregulation can be the result of both non-genetic and genetic insults that directly affect chromatin regulators.

Multiple sequencing studies have now revealed that epigenetic related factors are commonly mutated in cancer (Gonzalez-Perez, Jene-Sanz & Lopez-Bigas, 2013; Shen & Laird, 2013; Lawrence et al., 2014). Large-scale genomic analysis of point mutations in thousands of samples from distinct cancer types detected multiple epigenetic regulators as putative cancer driver genes (Lawrence et al., 2014). Interestingly, this analysis identified both shared epigenetic drivers but also ones that are unique to specific cancer types. Characteristic examples of universal drivers include factors with distinct epigenetic activities, like chromatin remodelling (*SMARCA4, ARID1B*), histone methyltransferase (*KMT2D, SETD2*) and histone acetyltransferase (EP300). On the other hand, mutations in DNA modifiers like *DNMT3A* and *TET2* were highly restricted to lymphocytic leukemia. Interestingly this analysis also revealed cases of epigenetic regulators (e.g. *EZH1, ARID2, MBD1*) that surpassed significance when examined across the pan-cancer cohort, agnostic of cancer type. This is in line with further data demonstrating that most of the epigenetic regulators are found mutated in low frequencies across cancer types (Gonzalez-Perez, Jene-Sanz & Lopez-Bigas, 2013).

Genomic studies have not revealed a preferential path towards epigenetic deregulation [Figure 4] (Brennan et al., 2013; Gonzalez-Perez, Jene-Sanz & Lopez-Bigas, 2013; Shen & Laird, 2013; Lawrence et al., 2014). On the contrary, mutations affect all known layers of epigenetic regulation (Shen & Laird, 2013), involving among others members of the SWI/SNF chromatin remodelling complex (St Pierre & Kadoch, 2017) and various writers, readers and erasers of histone modifications (Han et al., 2019). Recently, oncogenic mutations were also identified in histones which are the structural constituents of chromatin and act as the substrate of epigenetic regulation. These mutations often overlap with known histone modifications and are thought to hinder the function of the respective epigenetic writers and readers, adding another layer of complexity regarding the potential perturbations that can impact epigenetic function in cancer (Allis & Jenuwein, 2016). Interestingly, sequencing in glioblastoma patients revealed the existence of low-frequency mutually exclusive mutations

across multiple epigenetic regulators, further pointing to the fact that all layers within the epigenetic network get disrupted during carcinogenesis (Brennan et al., 2013).

Fitness altering mutations can in principle result in either gain or loss of the respective protein functionality. Gain of function mutations have been reported in a handful of chromatin regulators, with the most notable example being the frequent translocations affecting the H3K4 methyltransferase *KMT2A* that contribute to the pathogenesis in a subset of leukemias. However, analysis at the pan-cancer level revealed that the majority of the mutations in epigenetic factors lead to loss of their respective function [Figure 4], suggesting that epigenetic regulators act predominantly as tumor suppressors during carcinogenesis (Shen & Laird, 2013). The recurrent nature of these inactivating mutations along with the fact that the respective patients are often characterized by worse prognosis emphasizes their importance during cancer evolution (Kapur et al., 2013; Parker et al., 2016).

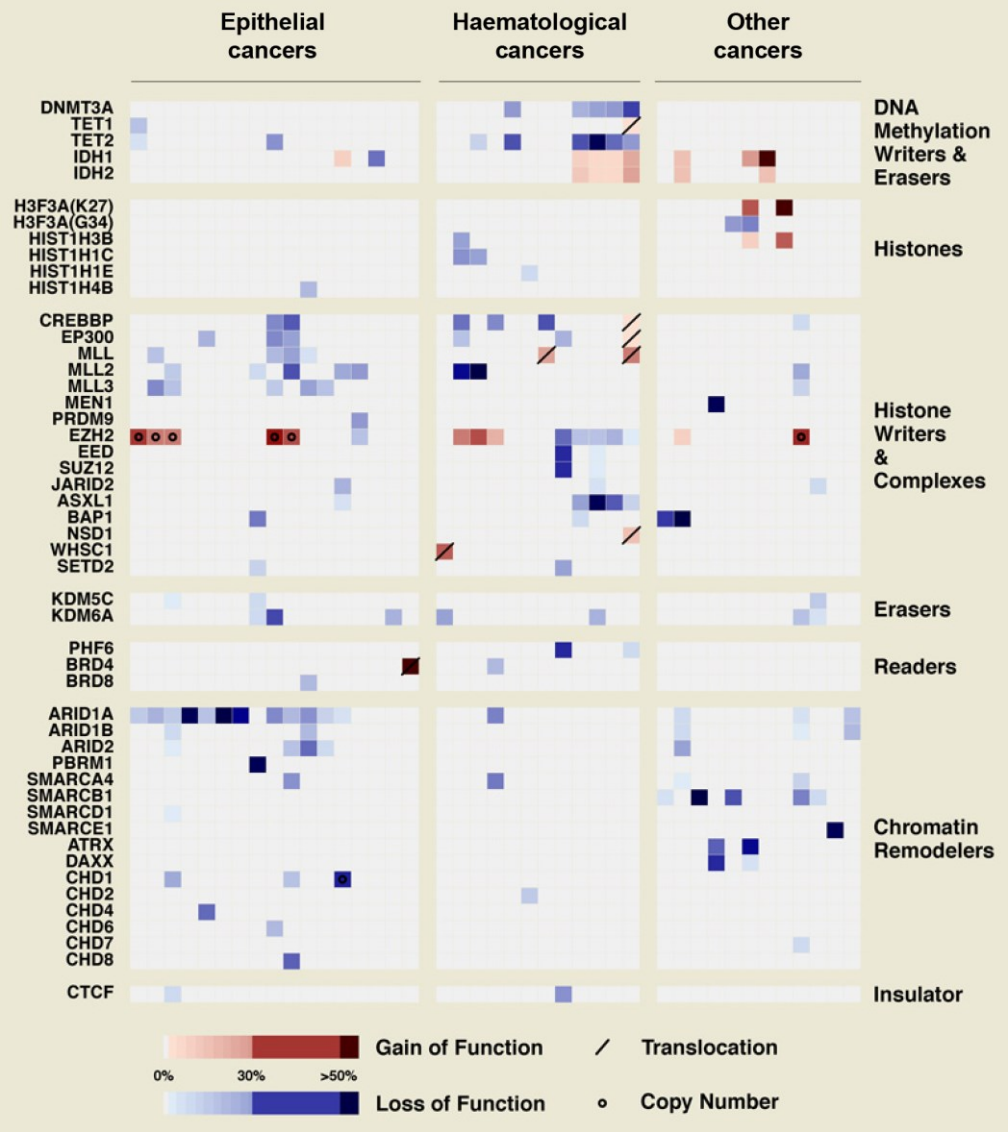


Figure 4: Recurrent inactivating mutations affecting diverse epigenetic regulators across multiple cancer types

Summary of genetic alterations across cancer types for a selected panel of epigenetic regulators belonging to distinct functional classes. The functional impact of each genetic event is classified as loss of function (blue) or gain of function (red). Translocation events are depicted with a dash while copy number alterations with a dot.

Figure adapted from Shen & Laird (2013).

Cancer type specific mutations in epigenetic regulators are frequent events and likely represent cases where disruption of the respective epigenetic activity promotes tumor initiation. One characteristic example comes from mutations in *DNMT3A* in hematological malignancies. DNMT3A and its paralogue DNMT3B are responsible for the *de novo* deposition of methyl marks on DNA, a modification associated with transcriptional repression (Allis & Jenuwein, 2016). In acute myeloid leukemia (AML) *DNMT3A* is recurrently affected by mutations and is associated with worse prognosis (Cancer Genome Atlas Research Network

et al., 2013). These events have been shown, via mechanisms that are dependent or independent of the catalytic activity, to inhibit the differentiation of hematopoietic stem cells leading to their expansion and initiation of tumorigenesis (Yang, Rau & Goodell, 2015). Interestingly *TET2*, which encodes for a protein that mediates DNA demethylation, is also frequently mutated in AML patients (Delhommeau et al., 2009) and promotes carcinogenesis by impairing the expression of myeloid differentiation genes (Tulstrup et al., 2021). This represents an interesting example where loss of function mutations on regulators that operate in antithetic manners can both result in a functional imbalance within an epigenetic layer (here DNA methylation) that is subsequently exploited to promote carcinogenesis. Analogous impairment of stemness differentiation programs have been also reported upon inactivation of epigenetic regulators in solid cancers, (Lewis et al., 2013; Concepcion et al., 2022) underscoring the importance of early mutations in diverse epigenetic regulators in promoting the acquisition of uncontrolled proliferation potential.

1.5.2 Epigenetic regulators are frequently mutated in cancer subclones

Inference of the temporal order in which mutations are acquired during cancer evolution, revealed frequent disruption in the epigenetic network across tumor subclones. Through sampling of spatially separated subpopulations in lung adenocarcinoma and squamous cell carcinoma, Hanjani and colleagues reported that subclonal mutations were enriched for specific functionally related genes like chromatin remodellers and histone modifiers along with factors involved in DNA damage response and signalling [Figure 5A] (Jamal-Hanjani et al., 2017). Accordingly, similar experimental approaches in renal carcinoma (Gerlinger et al., 2014, 2012) and glioma (Suzuki et al., 2015) identified subclonal mutations in genes associated with histone methylation (e.g. *SETD2, KDM5C*) and chromatin remodelling (e.g. *PBRM1, ARID1A, SMARCA4, SMARCC2, ARID2*). Finally, in silico inference of the mutational order from whole-genome sequencing data across 38 cancer types, identified enrichment of subclonal mutations in chromatin remodellers (e.g. *ARID1A, PBRM1*) and histone modifiers (e.g. *KMT2C/D*, and *SETD2*), suggesting that the late inactivating and selection of mutations in epigenetic regulators is a widespread phenomenon in cancer (Dentro et al., 2021). It is worth mentioning, that the current methodologies to reconstruct the timing hierarchy within the mutational landscape depend either on the sampling of a few spatially resolved samples or on the in-silico modelling from single biopsies, which both have limitations in uncovering subclonal diversity. Hence, the above observations are likely an underestimation of the actual prevalence of late mutations in epigenetic regulators.

One striking feature from the multi-region sequencing studies was the detection of parallel evolution[1], where spatially separated subclones within the same tumor core acquired distinct inactivating mutations in the same epigenetic regulator [Figure 5B] (Gerlinger et al., 2012, 2014; Suzuki et al., 2015). In glioma and renal cell carcinoma, this phenomenon extended also to cases where distinct populations carried mutations in different subunits (e.g. *SMARCC2, ARID2*) of the same functional complex (pBAF; chromatin remodelling) [Figure 5B] (Gerlinger et al., 2014; Suzuki et al., 2015). This is in line with previous observations from genomic studies showing that selection events can be better explained when epigenetic regulators are examined within their functional cohorts (Gonzalez-Perez, Jene-Sanz & Lopez-Bigas, 2013). Parallel evolution within distinct topological samples was also observed in methylation patterns in prostate cancer, further strengthening the connection between loss of strict epigenetic regulation and subclonal expansion (Brocks et al., 2014). Of note, the above studies examined both primary untreated and treated tumors, arguing that the observed patterns are not a result of therapeutic intervention but rather directly extend to the phenomena shaping tumor maintenance.

[1] In evolutionary terms, parallel evolution refers to the acquisition of similar traits in independent biological entities that co-exist within the same ecosystem.

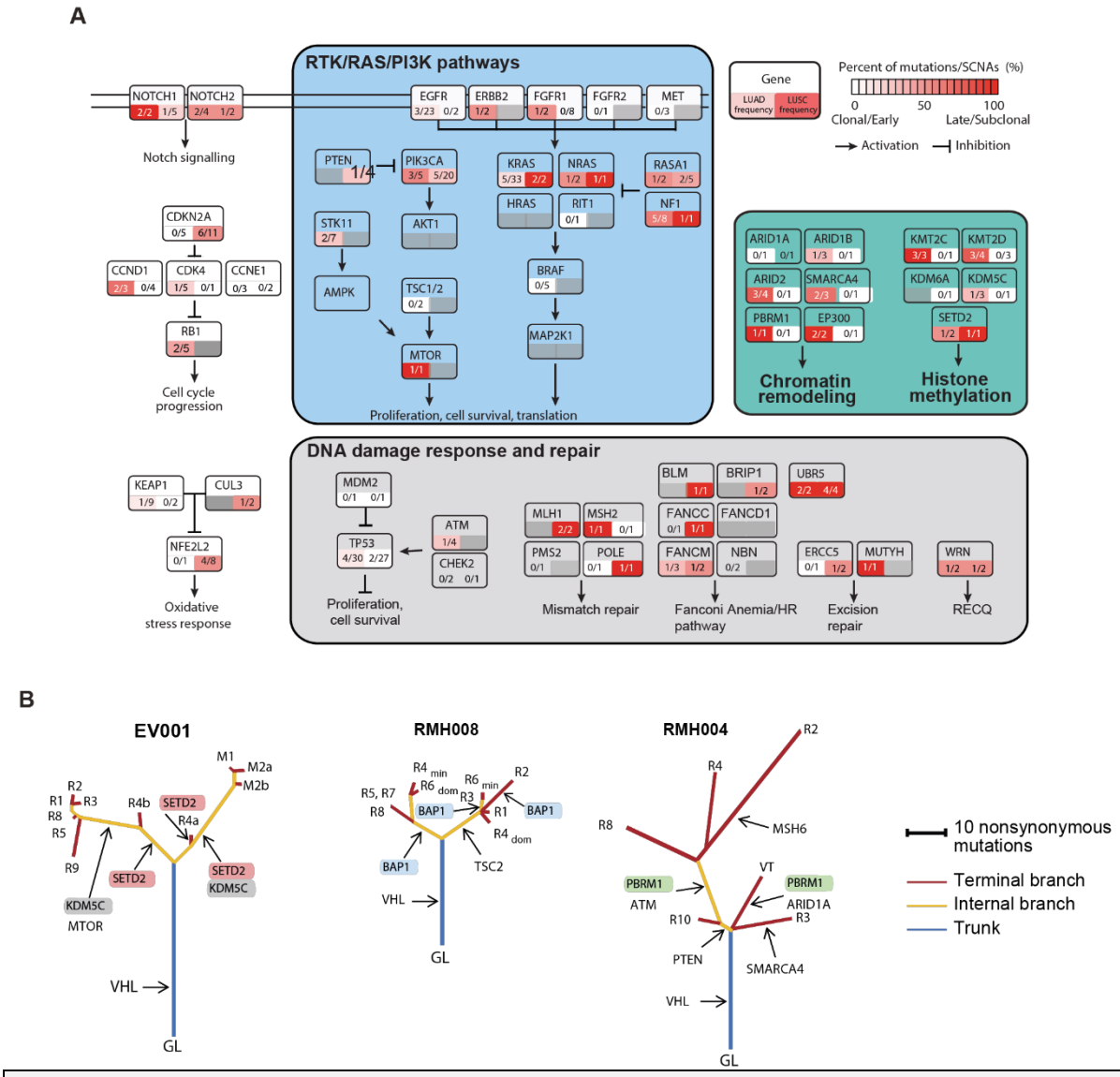


Figure 5: Subclonal inactivating mutations in core epigenetic regulators

[A] Timing of driver alterations during lung cancer evolution (TRACERx cohort). Data are presented for two cancer types, lung adenocarcinoma (LUAD) and squamous cell carcinoma (LUSC). The color within the box describes the relative abundance of subclonal to clonal genetic alterations for each gene.
[B] Phylogenetic trees of selected Clear Cell Renal Cell Carcinoma (ccRCC) tumors. Note the presence of parallel evolution events, where spatially distinct subpopulations within the tumor bulk acquire inactivating mutations in the same epigenetic regulators (EV001, RMH008) or across regulators functioning within the same protein complex (RMH004).

Figures adapted from Jamal-Hanjani et al (2017) and Gerlinger et al (2014).

The recurrent nature of epigenetic deregulation in subclones across various cancer types, suggests that epigenetic deregulation results in a phenotypic trait that is strongly favorable and thus selected over time. In contrast to the functional impact of early clonal mutations in epigenetic regulators, the significance of these late events remains elusive. Understanding the underlying mechanisms of these phenomena is crucial, considering that subclone

diversification is associated with worse prognosis in cancer patients (Landau et al., 2013; Jamal-Hanjani et al., 2017).

1.5.3 Functional impact of a deregulated epigenetic network during tumor maintenance

Examination of the mutational landscape within the epigenetic regulatory network revealed three important characteristics. First, in most of the cases mutations in epigenetic regulators result in loss of function of the respective protein. Second, the disruption is ubiquitous affecting similarly all known layers of epigenetic regulation. Third, during the later stages of tumorigenesis subclones preferentially acquire mutations in epigenetic regulators and even present extreme cases of parallel evolution.

What dictates this universal selective advantage of cancer cells that have lost their strict epigenetic control? One likely scenario is that inactivation of every single gene results in a different tumor-promoting cellular trait. From a systems perspective though, an intriguing alternative is that multiple different insults within the epigenetic network may converge to a similar phenotypic trait that is strongly favorable in cells during the later stages of tumorigenesis and thus is selected over time. Of note, these scenarios can also co-exist as epigenetic regulators can exert shared, overlapping and distinct functions within a cell. In the rest of this subchapter, I discuss several network-level effects that could emerge upon epigenetic deregulation and present data in support of either case.

Epigenetic deregulation and proliferation

Stemming from the definition of fitness one likely scenario is that disruption of diverse epigenetic regulators leads to an increase in the proliferation capacity of cancer cells, independent of the microenvironment, thus accounting for their expansion within subclones. Genome-wide sequencing studies have now revealed that epigenetic regulators act predominantly as tumor suppressors (Shen & Laird, 2013). Interestingly, various well characterised tumor suppressor genes, like *CDKN2A*, *TP53* and *PTEN,* that exert inhibitory effects on cell-cycle progression are often found subclonally mutated in cancer (Jamal-Hanjani et al., 2017; Dentro et al., 2021). In hepatocellular carcinoma, inactivation of the chromatin remodeller subunit *ARID2* resulted in derepression of cyclin D1 and cyclin E1, faster G1/S transitions and subsequently increased proliferation (Duan et al., 2016). Similar bypassing of G1/S checkpoint was also detected in cancer cell lines, upon inactivation of the histone acetyltransferase EP300 (Iyer et al., 2007). Finally, truncating mutations in *KMT2C*, a

methyltransferase frequently lost in subclones, resulted in aggressive tumors associated with increased expression of MYC targets and a concordant downregulation of cyclin-dependent kinase inhibitors (Limberger et al., 2022). Collectively, inactivation of various epigenetic regulators has been associated with increased cell-cycle progression, prompting towards further investigation.

Epigenetic deregulation and genetic instability

During the later stages of tumorigenesis, as discussed in 1.4, it is likely that fitness is strongly dependent on the ability of the cells to respond and survive to the diverse cues and selective pressures within their environment. Genetic diversity fuels disease progression, evident by the worse prognosis of patients with increased genetic diversity (Burrell et al., 2013; McGranahan & Swanton, 2017). Building on that, the identification of recurrent late mutations in genes involved in DNA damage response and repair, suggests that the resulting genomic instability is a favorable trait during subclonal expansion (Jamal-Hanjani et al., 2017).

Epigenetic regulators establish functional and structural alterations within chromatin and are involved in various processes including replication, transcription and DNA damage response. Thus, from a conceptual perspective epigenetic mechanisms can impact genome integrity, by altering both the frequency by which genetic abnormalities emerge (e.g. suppression of retrotransposons, chromosome segregation) and the fidelity of their resolution (e.g. DNA damage response) (Shen & Laird, 2013). Indeed, it is now well established that the genetic and epigenetic layers co-exist and significantly interact with each other to shape cellular behaviors during tumorigenesis (Shen & Laird, 2013). In line with that, inference of evolutionary histories from DNA methylation and mutational landscapes revealed highly similar patterns in both liquid and solid tumors, suggesting a co-dependency of genetic and non-genetic alterations during disease progression (Brocks et al., 2014; Oakes et al., 2014; Mazor et al., 2015).

An intriguing observation is that epigenetic regulators that were found to be frequently inactivated in cancer subclones have been linked with the emergence of genetic abnormalities. For instance, in renal carcinoma loss of *SETD2* that catalyzes the deposition of the repressive H3K36me3 mark, was associated with alterations in nucleosome occupancy followed by increased replication stress and impaired repair of the resulting DNA lesions (Kanu et al., 2015). Accordingly, cancer cells lacking *KMT2D*, writer of the “active” H3K4me3 mark, were characterised by increased transcriptional stress and mutational burden (Kantidakis et al., 2016). These studies provide initial evidence that epigenetic regulators that are often

subclonally mutated can lead to genomic instability. However, these studies focus on single genes, thus extrapolating towards the whole epigenetic network needs caution.

Epigenetic deregulation and transcriptional heterogeneity

Another favorable trait upon which epigenetic disruption may converge is increased transcriptional variability or phenotypic plasticity. These considerations arise from the inherent function of epigenetic mechanisms in shaping the transcriptional output of genes and subsequently cellular identity. As discussed earlier, in the literature the terms plasticity and heterogeneity are highly overlapping and in this subchapter, they will be discussed as such. During the experimental procedures in this thesis, further distinctions are made (Chapter 4). Transcriptional heterogeneity is important for cancer evolution (Fennell et al., 2022), metastasis (Nguyen et al., 2016; Quinn et al., 2021) and resistance to therapy (Sharma et al., 2010; Shaffer et al., 2017). The fact that splicing factors are recurrently mutated in subclones, provides a first line of evidence from the clinic regarding the importance of transcriptomic variability during subclone expansion (Dentro et al., 2021). Accordingly, in their recent study Fennell and colleagues observed that within competing clones, the ones demonstrating a selective advantage were characterized by increased transcriptional variance, further strengthening the notion that this trait is important for subclonal diversification (Fennell et al., 2022).

The identification of transcriptional heterogeneity in the absence of genetic drivers, pinpoints to an epigenetic underpinning. In line with that, single cell profiling studies have uncovered an association between DNA methylation (Johnson et al., 2021), chromatin accessibility (Corces et al., 2016) and discordance among different epigenetic modifications (Pastore et al., 2019) with transcriptional variability among individual cancer cells. Interestingly, Johnson and colleagues through multi-omics profiling of individual glioma cells observed significant disorder in the DNA methylation in regulatory elements of genes associated with cell identity and stress response, thus proposing an epigenetic underpinning of transcriptional variance that can be useful during tumor maintenance for example by aiding adaptation to stress (Johnson et al., 2021). However, the majority of these studies are descriptive, thus failing to uncover causal links within the examined properties of the system. In particular, without functional perturbation, it remains unclear if epigenetic disorder within cancer cells is a driver of transcriptional heterogeneity and if this is relevant for adaptation to hostile settings during subclone expansion. Another limitation of such correlative studies stems from the nature of

the epigenetic marks themselves, as their deposition within DNA and histones can be the result of transcriptional activity and not the causal driver of it.

In line with the need for functional interrogation, various studies within the field of therapy resistance have offered some initial insight into the causal contribution of epigenetic mechanisms towards favorable diversity. For instance, genetic or pharmacological inactivation of KDM5A/B methyltransferase altered the transcriptional variability within breast cancer cells, inhibiting their ability to adapt to therapeutic intervention (Hinohara et al., 2019). As discussed in 1.3.3.3 abrogation of other epigenetic regulators in diverse cancer models can shape the survival to therapeutic intervention that has been associated with transcriptional variability (Sharma et al., 2010; Ravindran Menon et al., 2015; Liau et al., 2017; Emran et al., 2018; Puig et al., 2018; Marsolier et al., 2022). It is worth noting that the perturbation of single epigenetic regulators fails to provide an overview regarding the importance of the epigenetic regulatory network to these processes. In their elegant study, Torre et al perturbed various regulatory networks (including the epigenetic layer), aiming to identify genes that can tune the emergence of transcriptionally distinct “primed” cells that drive resistance to targeted therapy (Torre et al., 2021). To do so, they focused on a well characterised model of melanoma adaptation to BRAF inhibition (Shaffer et al., 2017) and performed a pooled CRISPR screen targeting genes belonging to various regulatory layers (Torre et al., 2021). Their dual readout consisted of assessing both transcriptional variability and the number of resistant colonies, thus allowing for an observational dichotomy between transcriptional variability and actual resistance to stress. Targeting of various transcription factors, kinases and epigenetic regulators altered the resistance phenotype in melanoma cells, without a significant bias on fitness directionality or enrichment in the examined gene sets. Focusing on epigenetic regulators, revealed a handful of KO populations belonging to functionally distinct classes (histone modifiers and readers) that increased stress resistance. The authors also detected cases (e.g. cells KO for *KMT2D* and *KMT5B* methyltransferases) that exhibited increased resistance to therapy in the absence of any alteration in the number of primed cells. This raises the intriguing possibility that mutations in epigenetic regulators can favor survival via other mechanisms downstream of transcriptional priming that could be linked to the actual response of cancer cells to stress (Torre et al., 2021).

Epigenetic deregulation and response to stress

As stated earlier the conditions within the tumor microenvironment create a hostile setting that challenges cell fitness and survival. Despite the variability in the environmental conditions (Bhandari et al., 2019) and the fundamentally distinct properties that describe diverse cancer types, stress is considered a major component of tumorigenesis as reflected by the recurrent emergence of a stressed transcriptional state across cancer types (Baron et al., 2020; Barkley et al., 2022). To cope with potential stressors, cells have evolved complex systems that perform the following three functions: sensing unfavorable conditions, transmitting this information through signalling pathways and executing effector processes that promote adaptive behaviors to counteract stress (de Nadal, Ammerer & Posas, 2011). In response to diverse stimuli, cells can exhibit both common and stress-specific phenotypic alterations. In the majority of the cases, these stress responses converge to translational and transcriptional reprogramming in cells in order to halt the expression of growth-related genes and concomitantly favor the upregulation of stress-relates signatures (López-Maury, Marguerat & Bähler, 2008). This reprogramming allows a) the conservation of biomass until the stressor is resolved and b) the direction of resources towards transcriptional rewiring that can mediate adaptive behaviors. It is worth mentioning that the activation of the stress response can have both pro-survival and pro-death effects, mainly dependent on the severity of the stressor. Upon prolonged or harsh hostile conditions, the cells surpass a stress threshold and commit to apoptosis, a process mainly driven by the members of p53 family (Pflaum, Schlosser & Müller, 2014).

Although factors like cellular context, nature and duration of stress can elicit overlapping cell responses, it is well appreciated that the global downregulation of fitness genes is a point of convergence (López-Maury, Marguerat & Bähler, 2008). Indeed, this phenomenon has been observed in response to various stressors including nutrient deprivation, heat-shock, UV irradiation and therapeutic intervention (Emran et al., 2018; Gameiro & Struhl, 2018; Tufegdžić Vidaković et al., 2020; Cugusi et al., 2022). Considering the role of epigenetic regulators as mediators of environmental response and the fact that stress elicits changes at the transcriptomic level, a plausible scenario is that mutations in epigenetic regulators can hinder this process. Under harsh conditions, such inability may lower the probability of apoptotic induction thus providing a temporal competitive advantage to cells. It is worth noting that inactivating mutations in signalling pathways (e.g. MTOR) along with genes involved in oxidative stress response are enriched in subclones, suggesting that the inability of cells to respond to various internal and external stressors may be a beneficial trait during disease progression (Gerlinger et al., 2014; Suzuki et al., 2015; Jamal-Hanjani et al., 2017). Several lines of evidence from lower eucaryotes have demonstrated a strong contribution of epigenetic

regulators in stress response, involving alterations in chromatin occupancy at the proximal elements of stress genes, chromatin remodelling at gene bodies, modifications at various histone tail residues and interaction between stress-responsive transcription factors and chromatin modifiers (de Nadal, Ammerer & Posas, 2011). Overall, the widespread transcriptomic alterations that take place under stress and the implication of chromatin remodellers and modifiers in such responses, necessitate the investigation of a potential relationship between broad epigenetic deregulation and the ability of cancer cells to withstand various stressors within their niche.

Conclusion

The above alternate scenarios that could explain the selective advantage of epigenetically disrupted cancer cells are not mutually exclusive. Cancer is a complex evolving ecosystem where diverse favorable events can fuel disease progression at different time scales. This is partly attributed to the heterogeneous nature of the selective pressures imposed upon cancer cells that can range from transient environmental alterations to sustained chronic stress or therapy. For example, if the kinetic of the generation of *de novo* mutations is much smaller compared to the stressful environmental fluctuations it is expected that alternate mechanisms (like adaptation or resistance to stress) may drive events at the time. On the other hand, the gradual increase in genetic abnormalities may generate favorable diversity that can prove beneficial at later selective bottlenecks like metastatic dissemination and disease relapse.

Overall, during the later stages of cancer evolution disruption of the epigenetic network is a recurrent event that is strongly favorable and selected over time. Stemming from their ability to orchestrate events that are mediated through the chromatin (Transcription, DNA damage response, response to environmental stimuli, etc), alternate mechanisms could underpin the selective advantage upon their loss [Figure 6]. Large-scale correlative studies are limited in providing a causal link between the examined cellular properties. On top of that, most of the functional interrogations have so far focused primarily on single genes, thus providing limited insight into potential *network-level* effects upon epigenetic disruption. Further studies are required to explore in depth the functional consequence of disrupting the epigenetic network in various aspects of cancer evolution like survival in fluctuating stressful environments or resistance to therapy.

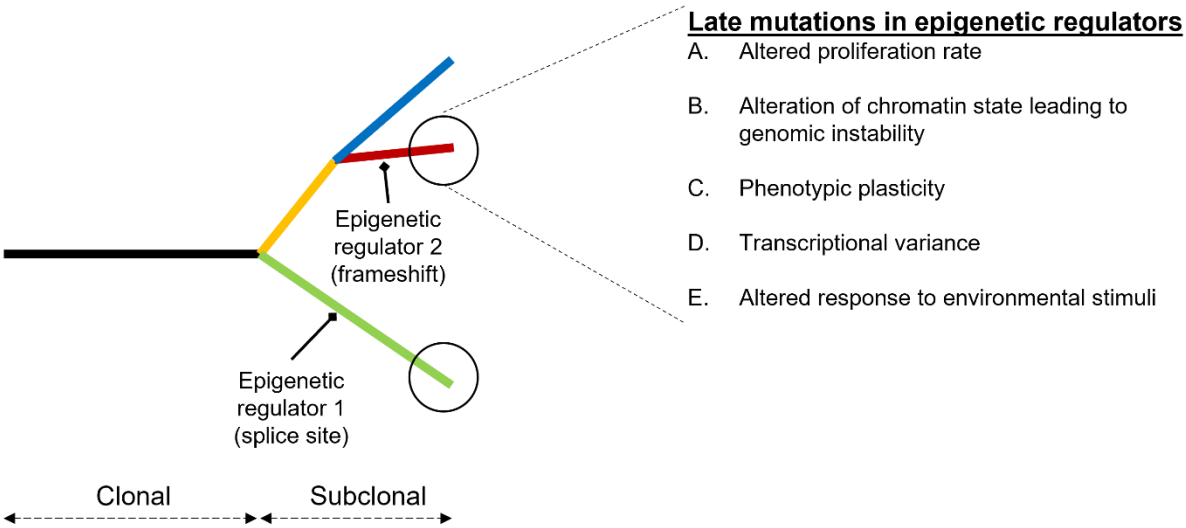


Figure 6: Potential network-level effects as a result of the widespread inactivation of epigenetic regulators during the later stages of tumorigenesis

As a network-level effect could be considered cases where perturbations in multiple and diverse nodes within the epigenetic network would converge to the same phenotypic outcome.

1.6 Synthesis of key introductory points and project aims

A dense synthesis of the key observations and concepts presented in the introductory sub-chapters can be found below:

Cancer is an evolutionary process governed primarily by selective forces that act on population heterogeneity to promote the propagation of the fittest. Heterogeneity arising from multiple different genetic and non-genetic mechanisms is pivotal for disease progression providing a plethora of diverging phenotypic traits upon which selection can act. Multiple lines of evidence suggest that the selectable unit is not gene mutations per se but rather the phenotypic traits that these events confer to malignant cells. Fitness is a quantitative trait that is not fixed but rather dynamic and is dependent on the context of cancer cells, namely the tumor-microenvironment (TME). TME consists of cellular and acellular components, is temporally and spatially heterogeneous and constantly poses new challenges to cells.

Epigenetic proteins are recurrently affected by inactivating mutations in cancer. Aside from high-frequency clonal mutations that represent initiating events during carcinogenesis, disruption of the epigenetic network is a phenomenon that is frequent in cancer subpopulations across various cancer types. This deregulation is broad, affects all known classes of epigenetic regulation and can even manifest in extreme cases of parallel evolution where the same regulator is affected by distinct inactivating mutations within the same tumor. Besides gene-level effects, these observations raise the intriguing possibility that diverse insults within the epigenetic network may converge to a common molecular trait that is favorable and thus

selected over time. Upon epigenetic deregulation, such a trait could be increased heterogeneity at the genomic or transcriptomic level or altered response to stressful environmental stimuli. Although there is initial evidence in favor of all the above scenarios, most of the current studies are either correlative or are limited to single perturbations within the network, thus failing to uncover general causal patterns.

Based on the above observations the aim of the current thesis is two-fold [Figure 7]:

i. Systematically inactivate multiple core epigenetic regulators and assess if this perturbation increases the survival of cancer cells in stressful environmental conditions, reminiscent of their tumor microenvironment [Chapter 3].

ii. If this is the case, then explore various molecular traits that could underpin the phenotypic deviation conferred by the broad disruption of the epigenetic network [Chapter 4].

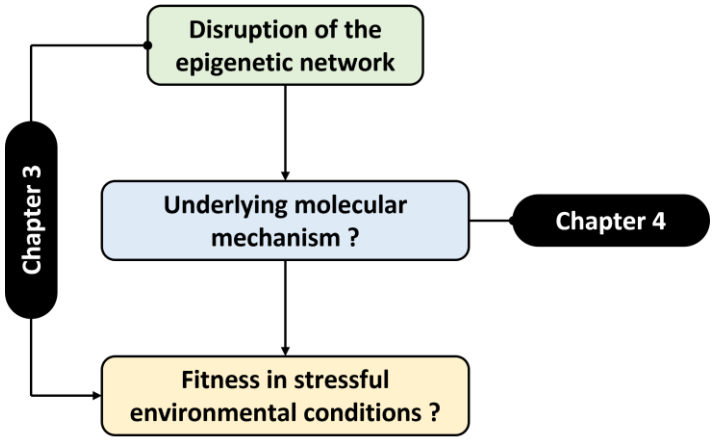


Figure 7: Schematic representation of the conceptual approach followed in this PhD thesis

The aim is to investigate the effect of epigenetic deregulation in the fitness of cancer cells under environmental stress [Chapter 3] and to explore potential molecular mechanisms that could underly such a dependency [Chapter 4]

Chapter 2. Materials and Methods

Note: Several parts of this chapter are already published (Loukas et al., 2023; Simeoni et al., 2023). *Selective advantage of epigenetically disrupted cancer cells via phenotypic inertia* © 2023 by Ioannis Loukas et al is licensed under CC BY 4.0. *CRISPR-based large-scale modeling of loss-of-function mutations to investigate mechanisms of stress resistance in cancer* © 2023 by Fabrizio Simeoni is licensed under CC BY 4.0. The researchers that contributed to this chapter are acknowledged accordingly at the beginning of the respective subchapters.

2.1 Experimental models

2.1.1 Cellular models and culture conditions

All the cell lines used in this study are listed in Table 2. Parental lines and their *de novo* generated derivatives were cultured in RPMI 1640 (ThermoFisher Scientific, #52400025) supplemented with 10% fetal bovine serum (FBS), 100 U/ml penicillin, and 100 µg/ml streptomycin at 37 °C in 5% $C0_2$. Patient-derived xenografts MEXF 2090 and LXFL 1674 were obtained from the Charles Rivers tumor model compendium (for their characteristics see Table 3). To enable efficient CRISPR-Cas9 mediated KO of the desired genes under study, cell lines expressing an inducible form of Cas9 were generated by a former member of the Scaffidi lab (Louise Richardson, MSc thesis, 2017). Briefly, MEXF 2090 and LXFL 1674 cells were transduced with a lentiviral pCW-Cas9 vector (Wang et al., 2014; Monserrat et al., 2021). Following a 7-day selection with 600 µg/mL hygromycin B (ThermoFisher Scientific, #10687010), clones with minimal background levels of Cas9 and responsive to induction with 1 µg/mL doxycycline (Sigma-Aldrich, #D9891) were isolated. The clonal lines derived from PDX MEXF 2090 and PDX LXFL 1674 were named PDX MeA5a and L1C5c respectively. Throughout this thesis the initial parental names are used.

2.1.2 Mice models

Male NSG Mice (NOD.Cg-Prkdc^{scid}Il2rg^{tm1Wjl}/SzJ) used in this study for the *in vivo* competition experiment were obtained from the common Crick colony. Mice were used at 11-13 weeks of age for the experiment. The mice were housed in well-ventilated cages at a constant temperature and humidity (23 °C ± 2 °C, 50–60%) in a pathogen-free controlled environment, with a standard 12 h-12 h light-dark cycle, and unrestricted access to water and food. The weight and signs of diseases were monitored at regular intervals.

All the *in vivo* experiments were approved by the Francis Crick Institute's Animal Welfare and Ethical Review Body and conformed to UK Home Office regulations under the Animals (Scientific Procedures) Act 1986 including Amendment Regulations 2012. For subcutaneous transplantation experiments, the maximum tumor size did not exceed a mean diameter of 1.2 cm, in compliance with project license 70/8931 and PP9490916.

2.2 DNA and RNA analysis

2.2.1 Plasmid extraction

A list of all the plasmids used in this study (purchased and *de novo* generated) along with a brief description of their applications and growth conditions is presented in Table 4. For plasmid extraction, depending on the desired scale, mini or maxi bacterial preparations with the respective antibiotic were generated followed by isolations using QIAGEN® Plasmid Mini Kit (#12123) or QIAGEN® Plasmid Maxi kit (#12165) as per the manufacturer's instructions.

For the lentiviral packaging plasmids, an additional purification step was performed, via phenol/chloroform purification. Briefly, the eluted plasmids were mixed with an equal volume of phenol:chloroform:isoamyl alcohol (25:24:1, v/v) and vortexed thoroughly for 30 sec. After centrifugation at 18000 g for 5 min, the upper aqueous phase was extracted and mixed with 1/10 volume of 3M Sodium acetate (CH_3COONa) and three volumes of 100% ethanol. After incubation of the mixture for 30 min at -80 °C, the plasmid was precipitated by centrifugation at 18000 g for 15 min at 4 °C. Finally, the pellet was washed with 70% ethanol, to achieve salt removal, dried at room temperature and resuspended in nuclease-free H_2O.

The concentration and purity of the isolated plasmid were assessed by measuring the 260/280 and 260/230 ratios in a NanoDrop™ 2000/2000c Spectrophotometer (ThermoFisher Scientific, #ND-2000). All isolated plasmids were stored at -20 °C. To ensure long-term maintenance for all plasmids, also glycerol stocks (25% v/v glycerol) were generated and maintained at -80 °C.

2.2.2 Genomic DNA extraction (gDNA)

For gDNA isolation two distinct methodologies were applied depending on the format and scale of the respective experiment. gDNA extraction was performed using the DNeasy Blood & Tissue kit (Qiagen, #69506) following the standard manufacturer's protocol. For gDNA isolation for tumor xenografts refer to section 2.13. For high-throughput gDNA isolation from 96-well plates (like the one performed for exome-sequencing) a kit-free protocol was applied

by using the Bradley lysis buffer followed by ethanol precipitation. For the production of the lysis buffer see the table below.

Bradley Lysis Buffer		
Stock Solutions	**Final Concentration**	**For 500mL**
1M Tris-HCl (pH 7.5)	10mM	5 mL
0.5M EDTA	10 mM	10 mL
10% SDS	0.5%	25 mL
5M NaCl	10 mM	1 mL
H_2O	N/A	bring up to 500ml

Briefly, MEXF 2090 and LXFL 1674 cells grown in 96-well plates were washed once with PBS, lysed with 50 µl Bradly Lysis buffer supplemented with 1 mg/ml Proteinase K (Qiagen, #19131) and incubated overnight at 60 °C. DNA was precipitated by the addition of 100 µl of ice-cold EtOH/NaCl (75 mM NaCl in 100% EtOH) followed by vigorous mixing, incubation at room temperature for 30 min and centrifugation at 3000 g for 20 min. The pellet was washed twice with 100 µl of 70% EtOH followed by centrifugation at 3000 x g for 10 min. DNA was resuspended in 30 µl of warm TE pH 8.0 after incubation at 56 °C for 10 min. Sample concentration and purity were assessed by measuring the 260/280 and 260/230 ratios in a NanoDrop™ 2000/2000c Spectrophotometer (ThermoFisher Scientific, #ND-2000).

2.2.3 Polymerase Chain Reaction (PCR)

All the PCR reactions described in this study were performed using the high-fidelity Herculase II Fusion Polymerase (Agilent Technologies, #600679). All the primers used in this study can be found in Table 5.

For the PCR amplification of mCherry-NLS from pUAS-mCherry-NLS (Addgene, #87695) the following experimental conditions were used (see also Methods section 2.3.1).

Reagent	Quantity per rxn	Temperature and Cycles
Plasmid vector	25 ng	95 °C for 2 min (activation)
dNTP mix (25 mM each dNTP)	0.5 µl	95 °C for 30 sec (denaturation)
Herculase II reaction buffer (5X)	10 µl	55-69 °C for 30 sec (annealing)
Primer mix (10 µM each)	1.25 µl	72 °C for 30 sec (extension)
Herculase II fusion DNA polymerase	0.5 µl	Return to denaturation, 35X
DMSO	2 µl (4% final conc)	72 °C for 3 min (extension)
H_2O	Up to 50 µl	Hold at 4 °C

For the PCR amplification of sgRNA loci from genomic DNA the following nested PCR set-up was applied (see also Methods section 2.13).

Step 1 of 2 for nested PCR (PCR1)

Reagent	Quantity per rxn	Temperature and Cycles
Genomic DNA	2 µg	98 °C for 3 min (activation)
dNTP mix (25 mM each dNTP)	1 µl	98 °C for 20 sec (denaturation)
Herculase II reaction buffer (5X)	20 µl	60 °C for 20 sec (annealing)
Primer mix (10 µM each)	1.25.µl	72 °C for 30 sec (extension)
Herculase II fusion DNA polymerase	1 µl	Return to denaturation, **20X**
DMSO	4 µl (4% final conc)	72 °C for 3 min (extension)
H_2O	Up to 100 µl	Hold at 4 °C

Step 2 of 2 for nested PCR (PCR2)

Reagent	Quantity per rxn	Temperature and Cycles
PCR1 clean product (1/50 diluted)	10 µl	98 °C for 3 min (activation)
dNTP mix (25 mM each dNTP)	2 µl	98 °C for 20 sec (denaturation)
Herculase II reaction buffer (5X)	40 µl	58 °C for 20 sec (annealing)
Primer mix (10 µM each)	2.5.µl	72 °C for 30 sec (extension)
Herculase II fusion DNA polymerase	2 µl	Return to denaturation, **28[2]X**
DMSO	8 µl (4% final conc)	72 °C for 3 min (extension)
H_2O	Up to 200 µl	Hold at 4 °C

2.2.4 PCR amplicon purification

Before using PCR products for subsequent experiments (sequencing, digestion with restriction enzymes, etc), the amplicons were further purified using QIAquick PCR Purification Kit (#28106) as per the manufacturer's instructions. Briefly, five volumes of buffer PB (binding buffer) were mixed with one volume of PCR reaction. To bind DNA, the mixture was applied to a silica-based column and centrifuged for 60 sec at 18.000 g. The column was then washed with 750 µl buffer PE, followed by dry centrifugation and elution in nuclease-free H_2O.

[2] To define the least number of cycles that provide maximal amplicon output, a titration of PCR cycles was performed ranging from 7 to 31 cycles in 3-cycle increments. At cycle 28 the amplicon production started to plateau as indicated by the band intensity in an agarose gel.

2.2.5 Gel electrophoresis and extraction

Analysis of plasmids, RNA and PCR products was performed via agarose gel electrophoresis. Samples were mixed with 10x loading dye (ThermoFisher Scientific, #10816015) and run on gels composed of 0.6-2% agarose in 1X TBE (Tris-Borate-EDTA), supplemented with 1X ethidium bromide (Promega, # H5041). To estimate band size either Quick-Load Purple 50 bp DNA Ladder (NEB, #N0556S) or GeneRuler 1kb DNA ladder (ThermoFisher Scientific, #SM0312) were used depending on the analytical purpose. The samples were run under 2-4 V/cm voltage and after sufficient separation the electrophoretic pattern was visualized under exposure to UV light.

For the isolation of specific DNA fragments of interest (for example during cloning), the desired band was excised using a scalpel, under minimal UV illumination, and subsequently purified by using QIAquick Gel Extraction Kit (#28706) as per the manufacturer's guidelines. Sample concentration and purity were assessed vis spectroscopic analysis in NanoDrop™ 2000/2000c (ThermoFisher Scientific, #ND-2000).

2.2.6 Total RNA extraction

For transcriptomic analysis, total RNA was isolated from the indicated cell populations by using Qiagen RNeasy Plus Micro kit (#74034) as per the manufacturer's protocol. To achieve substantial starting material from experimental setups that were performed in 96-well plates, 4 to 6 replicate wells were pooled depending on cellular density. Briefly, lysis buffer (RLT) was added to a maximum of 500.000 cells, incubated for 5 min and then transferred through a column that binds genomic DNA. The flow-through was then mixed with an equal volume of 70% ethanol and loaded into the silica-based membrane for RNA binding. The column was subsequently washed 3 times with a sequence of respective buffers and 80% ethanol followed by elution in RNAse-free H_2O. Total RNA concentration and purity were estimated by measuring the 260/280 and 260/230 ratios in a NanoDrop™ 2000/2000c Spectrophotometer (ThermoFisher Scientific, #ND-2000). All samples were stored at -20 °C (short term; <1 week) or -80 °C (long-term storage).

2.2.7 Generation of complementary DNA (cDNA)

Generation of cDNA from total RNA was achieved by using a High-Capacity cDNA Reverse Transcription Kit (ThermoFisher Scientific, #4368814) following the manufacturer's instructions. In all the experiments described in this study 0.5 µg of total RNA was used as an input for cDNA synthesis. The quantities of the respective reagents and the reaction conditions in the C1000 Touch Thermal Cycler (Biorad, #1851148) are described in the table below.

Reagent	Quantity per rxn	Temperature
RNA	500 ng	25 °C for 10 min
10X RT buffer	2 µl	37 °C for 120 min
25X dNTP	0.8 µl	85 °C for 5 min
10X primers	2 µl	Hold at 4 °C
Enzyme	1 µl	
Nuclease-free H_20	Up to 20 µl	

2.2.8 Quantitative PCR (qPCR)

For quantification of transcript levels, cDNA from 0.5 µg of total RNA was diluted 1/10 and used as input for RT–qPCR using SsoAdvanced™ Universal SYBR® Green Supermix (Bio-Rad, #172-5274) on a CFX96 real-time PCR detection system (Bio-Rad). *PPIA* was used as a reference gene for normalization. Primers used for RT–qPCR in this study are listed in Table 5. The reaction conditions along with the cycling set-up can be found in the table below.

Reagent	Quantity per rxn	Temperature and Cycles
Universal Green Supermix (2X)	10 µl	95 °C for 30 sec (activation)
Primers (5 µM)	2 µl	98 °C for 10 sec (denaturation)
cDNA (1/10 from 500ng)	2 µl	60 °C for 30 sec (annealing/ extension/ plate reading)
Nuclease-free H_20	6 µl	Return to denaturation, **35X**
Total volume	20 µl	65 to 95 °C in 0.5 °C increments at 5 sec / step (Melt Curve analysis)

2.3 Molecular Cloning

2.3.1 Generation of pTRIP-SFFV-mCherry-NLS

To generate a plasmid vector expressing a nuclear restricted form of mCherry fluorophore (pTRIP-SFFV-mCherry-NLS), the following plasmids were purchased; pTRIP-SFFV-EGFP-NLS (Addgene, #86677) and pUAS-mCherry-NLS (Addgene, #87695) with the scope being to substitute EGFP for mCherry in the former vector. To do so, primers with overhangs carrying sequences for the restriction enzymes BamHI and XhoI were used to amplify mCherry-NLS from the pUAS vector. Subsequently, both the PCR amplicon and the destination vector (pTRIP-SFFV-EGFP-NLS) were digested by BamHI and XhoI in FastDigest Buffer (ThermoFisher Scientific, #FD0055, #FD0694 and #B64 respectively) for 1 hour at 37 °C. For the primer set used and the respective reaction conditions refer to Table 5 and section 2.2.3 respectively. The products of the above digestions were separated via electrophoresis in an 1% agarose gel and the desired fragments (insert and vector without EGFP) were excised and purified via gel extraction using QIAquick Gel Extraction Kit (Qiagen, #28706).

Based on their size, insert and vector were mixed in a 9:1 molecular ratio and ligated using the Quick Ligation™ Kit (NEB, #M2200S), as per the manufacturer's guidelines. After a 15 min incubation at room temperature, the ligation mixture was transformed into One Shot™ Stbl3™ Chemically Competent *E. coli* (ThermoFisher Scientific, #C737303). Briefly, the mixture was incubated on ice for 30 min, followed by a heat shock at 42 °C for 45 sec. Bacteria were recovered on ice for 2 min, cultured in S.O.C. media (NEB, #B9020S) for 1 hour at 30 °C and then plated in ampicillin containing Luria broth (LB) agar plates. After 24 h of growth at 30 °C, single colonies were picked, incubated overnight in LB media with 100 μg/mL Ampicillin and the plasmids were isolated via QIAGEN® Plasmid Mini Kit (#12123).

The successful generation of pTRIP-mCherry expressing constructs was verified by both test digestions with restriction enzymes and after transfection of HEK293T with the putative isolated plasmids. 48 h post transfection, cells were imaged in IncuCyte S3. The plasmid generating the best nuclear to cytoplasmic intensity ratio was selected to move forward.

2.3.2 Generation of pLenti-BSD-sgRNA-Capture_seq_1E

For the multiplexed single-cell transcriptomics experiment (refer also to sections 2.15.3 and 4.3.3.1), it was required to generate a pLenti-BSD-sgRNA vector with a modified sgRNA construct carrying a unique oligonucleotide sequence (*Capture Sequence 1*) that can be

recognized by the 10X Genomics pipeline[3]. To generate this vector, the following plasmids were used; pLenti-BSD-sgRNA and PUC57-Capture_sequence_1 (Synthetic sgRNA construct purchased from Genscript) with the scope being to substitute the conventional sgRNA scaffold for the modified one in the pLenti vector. Both plasmids were digested with BstBI (NEB, #R0519S) and NsiI (NEB, #R0127L) in NEBuffer 2.1 (NEB, #B6002SVIAL) for 1 hour at 37 °C and the products were analyzed in an agarose gel (0.6 and 2% for vector and insert respectively). The desired bands (insert from PUC57 and pLenti vector without sgRNA scaffold) were excised and purified via gel extraction using QIAquick Gel Extraction Kit (Qiagen, #28706).

Based on their size, insert and vector were mixed in a 9:1 molecular ratio and ligated using the Quick Ligation™ Kit (NEB, #M2200S), as per the manufacturer's guidelines. After a 15 min incubation at room temperature, the ligation mixture was transformed into One Shot® ccdB Survival™ 2 T1R Chemically Competent Cells (ThermoFisher Scientific, #A10460). Briefly, 10 µl from the mixture was incubated on ice for 30 min, followed by a heat shock at 30 °C for 42 sec. Bacteria were recovered on ice for 2 min, cultured in S.O.C. media (NEB, #B9020S) for 1 hour at 30 °C and then plated in ampicillin containing Luria broth (LB) agar plates. After 36 h of growth at 30 °C, single colonies were picked, incubated overnight in LB media with 100 µg/mL Ampicillin and 25 µg/mL chloramphenicol and the plasmids were isolated via QIAGEN® Plasmid Mini Kit (#12123).

The success of the cloning procedure was verified by the electrophoretic pattern after test digestions and via Sanger sequencing (Illumina outer 1R: CCTCGACCTGCTGGAATCTC) performed by the Genomics Equipment Park facility at the Francis Crick Institute.

2.3.3 Cloning sgRNAs into pLenti-BSD-sgRNA-Capture_seq_1E

For the generation of individual sgRNAs, primers containing the sgRNA sequence and a restriction site against BsmBI were hybridized in the presence of an annealing buffer (10 mM TRIS + 1 mM EDTA). Then the solution was incubated at 95 °C for 5 min followed by a reduction to 25 °C in increments of 0.1 °C / sec. The annealed oligos were then diluted 1/200 and 2 µl were used as input for cloning into the pLenti-BSD-sgRNA-Capture_seq_1E plasmid via Golden Gate (GG) assembly technology (NEB). The BsmBI restriction site is eliminated

[3] For additional details regarding the sgRNA scaffold refer to the 10X Genomics' technical note regarding the "Guide RNA Specifications Compatible with Feature Barcoding technology for CRISPR Screening"(https://cdn.10xgenomics.com/image/upload/v1660261286/support-documents/CG000197_GuideRNA_SpecificationsCompatible_withFeatureBarcodingtechnology_forCRISPRScreening_Rev-A.pdf)

from the ligated product, so digestion and ligation can be carried out simultaneously. The relative quantities of the reagents and the reaction conditions are listed below.

Reagent	Quantity per rxn	Temperature and Cycles
BsmBI (5 units)	0.5 µl	45 °C for 2 min (Digestion)
Annealed oligos (1.2ng)	2 µl	20 °C for 2 min (Ligation)
pLenti_BSD_sgRNA (50ng)	1 µl	Return to Digestion, 25X
T7 DNA Ligase (1500 units)	0.5 µl	60 °C for 10 min (Linearise residual vector)
10X PNK Buffer (+66µM ATP)	1 µl	80 °C for 10 min (Enzyme heat inactivation)
Nuclease-free H_20	Up to 10 µl	Hold at 4 °C

1 µl from the GG reaction was used as input for transformation of One Shot™ Stbl3™ Chemically Competent *E. coli* (ThermoFisher Scientific, #C737303). Briefly, the mixture was incubated on ice for 30 min, followed by a heat shock at 42 °C for 45 sec. Bacteria were recovered on ice for 5 min, inoculated in Luria broth (LB) supplemented with 100 µg/mL Ampicillin and grown overnight at 37 °C.

The successful cloning of the desired sgRNAs into pLenti-BSD-sgRNA-Capture_seq_1E was verified by the electrophoretic pattern after test digestions and via Sanger sequencing (Universal U6 promoter) performed by the Genomics Equipment Park facility at the Francis Crick Institute. The newly generated pLenti-BSD-sgRNA-Capture_seq_1E plasmids carrying the desired sgRNAs were then transduced into MEXF 2090 cells and KO of the targeted genes was induced as described in 2.5. The cutting efficiency was assessed a) directly by Sanger sequencing of the targeted loci and b) indirectly by assessing the phenotypic behavior of the KO populations under nutrient starvation. Quantitative assessment of genome editing by Tracking of Indels by DEcomposition (TIDE) analysis (Brinkman et al., 2018) revealed minimal activity across the majority of the tested single sgRNAs. In line with that, under nutrient starvation most of the KO populations demonstrated no or mild fitness advantage compared to control populations (in contrast to the detected behaviors in the large-scale fitness assays). Based on these observations I concluded that the newly generated plasmids were not functional. Notably, this inactivity was observed in single sgRNAs that were validated in other studies within the lab (Mortimer et al., 2019), thus indicating that the inactivity is not a result of selecting a non-functional sgRNA out of the library pool. Therefore, for the multiplexed scRNA-seq experiment, I used these non-functional sgRNAs to *barcode* MEXF 2090 cells that were already KO for the desired epigenetic regulators (also refer to section 2.15.3).

2.4 Production of lentiviral particles

To generate viral particles, a second-generation lentiviral production system was used consisting of the following three plasmids: 1) packaging plasmid (psPAX2; Addgene, #12260), 2) envelope expressing plasmid (pMD2.G, Addgene, #12259) and 3) transfer plasmid (pLenti or pTRIP vectors for expression of sgRNAs or fluorophores respectively). To achieve higher lentiviral titers, pAdvantage (Promega, #E1711) was also included in the aforementioned plasmid mixture. For viral production, 90% confluent HEK293T cells (Cell Services, Francis Crick Institute) were transfected with the plasmid mixture by using FuGENE® HD (Promega, #E2311), at a ratio of 3:1 FuGENE® HD to DNA. The viral productions were optimized and performed in large-scale in 96-well plates. Wherever other formats were used, the reagent quantities listed below were scaled accordingly based on the relative working surface area.

Transfection in a 96-well (1x rxn)	
Reagent	**Quantity**
pLenti-sgRNA or pTRIP vector (Transfer)	68 ng
psPax2 (packaging)	50.6 ng
pMD2.G (envelope)	16.8 ng
pAdvantage (Increased yield)	15 ng
Total DNA	150 ng
FuGENE HD	0.45 µl

24 and 48 h post transfection, the supernatant containing the viral particles was collected, pooled and filtered through a 0.45 µm filter plate (Millipore, #MSHVS4510) after centrifugation at 3000 g for 1 min. The collected virus was then stored at -80 °C. To define the best viral dilution for subsequent experiments, 5000 melanoma or 8000 lung cancer cells in 96-well plates were infected with a titration of GFP expressing virus in the presence of 5 µg/mL Polybrene (Santa Cruz, #sc-134220), used as a proxy to estimate the infection efficiency. In most of the cases the lowest viral titration that yielded maximal infection efficiency and no significant growth deficit was used.

2.5 Generation of KO cell lines

To generate CRISPR-Cas9 mediated KO populations, cells inducibly expressing Cas9 upon doxycycline induction (1 µg/mL) were transduced with lentiviral particles expressing individual sgRNA constructs (For viral production refer to section 2.4). To select for the infected cells, 48 h post transduction, blasticidin (6 µg/mL) was added to the media. The successful selection was confirmed by the elimination of replicate populations grown in parallel but not infected. The above procedure was performed in various scales (generation of individual KOs for fitness validations or KO libraries for large-scale fitness assays) and in various cellular models (MEXF 2090, LXFL 1674, MEXF 2090 EGFP-NLS, MEXF 2090 mCherry-NLS). The desired sgRNA plasmids were sourced from an available arrayed lentiviral sgRNA library (Henser-Brownhill, Monserrat & Scaffidi, 2017).

After selection, expansion and induction of KO cells for 10 days, individual populations were frozen in FBS supplemented with 10% DMSO and stored in liquid nitrogen. On the other hand, the libraries of the KO cells were frozen at -80 °C in multiple aliquots. To freeze populations in 96-well format, cells were detached from the plate using 30 µl trypsin per well. Following the addition of 80 µl of FBS containing 10% DMSO, plates were sealed and stored at -80 °C for up to 4 weeks. For long-term preservation copies of the libraries were thawed, propagated and frozen again. To thaw cells, plates were placed in a water bath at 37 °C for a few seconds and spun for 5 min at 4 °C after the addition of 50 µl of medium to each well. Fresh medium (100 µl) was finally added to each well after removal of 120 µl of freezing medium.

2.6 Large-scale and validation fitness assays

2.6.1 Plate layout

Also refer to Loukas et al (2023) and Simeoni et al (2023).

From the available sgRNA library (Henser-Brownhill, Monserrat & Scaffidi, 2017), constructs targeting 318 genes, encoding core epigenetic regulators were selected and arranged in eight 96-well plates (Corning, #3596). For the full list of the targeted regulators and their classification to respective epigenetic families refer to Table 1 and Table 8. Each plate consisted of 40 KO populations in which distinct epigenetic regulators were inactivated, along with 20 negative control populations transduced with sgRNAs targeting 5 non-expressed genes, four replicates each. Multiple negative controls per plate were included to account for technical variability and well-effects, and to enable robust normalization of the observed phenotypes across plates and experiments. Each plate contained a well with *ARID2*-targeting

sgRNAs as an inter-plate standard. External wells were excluded and filled with PBS to avoid edge effects.

2.6.2 Assay pipeline

During the plating step for the large-scale fitness assays, multiple identical replicates, sufficient for the different treatments, were generated from each library plate. To do so, in each plate the medium was discarded and cells were washed with PBS followed by addition of 30 µl of trypsin and incubation at 37 °C for 5 min. Then, 170 µl were added in each well and cells were resuspended by vigorous pipetting. The cell suspension transferred to a 2 ml deep well block and mixed with a specific volume of RPMI medium to achieve the desired dilution. Finally, 50 µl per population were seeded into the 96-well plates, already containing 50 µl of fresh medium. The plates were shacked to achieve homogeneous plating and then incubated at 37 °C. Approximately 24 hr post seeding, when the median cell count/well across plates reached ~4000 cells, the plates were either grown under stressful conditions (see below) or maintained in unperturbed conditions. Over the course of the experiment one representative plate was monitored by time-lapse imaging to confirm that the growth kinetics were as expected based on pilot experiments. At the indicated endpoints for each condition (see below) the population fitness was assessed by quantification of cell count. To do so, cells were fixed with 4% paraformaldehyde (PFA, Alfa Aesar, #43368) followed by permeabilization with 0.5% Triton X-100 in PBS and nuclei staining with SYTOX™ Green Nucleic Acid Stain (ThermoFisher Scientific, #S7020). Imaging and quantification were performed using an Incucyte® S3 Live-Cell Analysis System.

2.6.3 Stress conditions

KO and control populations were cultured in the following conditions: a) unperturbed b) glutamine starvation c) acidic environment and d) replication stress, with each stress applied at two distinct strengths. Glutamine deprivation was sustained for 3 or 7 days after glutamine removal. Media acidification was induced by addition of HCl to a final pH of 6.7 or 6.5, which resulted in a 40-60% reduction in cell counts after 2 days compared to untreated cells. To induce replicative stress, cells were cultured in the presence of 200 µM or 250 µM hydroxyurea (Sigma-Aldrich, #H8627) which resulted in 40% or 60% in cell counts after 2 days compared to untreated cells. Cell count for populations grown in unperturbed conditions was quantified at day 2. The endpoint of each treatment was determined by when the fittest population reached confluence (to maintain linearity of comparisons across populations) and depended on how severely cells were affected by each stress, with deprivation of L-glutamine being the

most deleterious stress. A survival benefit in at least one strength condition for each stress was considered as enhanced fitness.

2.6.4 Data filtering

Before quantifying stress-specific fitness, various parameters were assessed to remove: 1) wells containing cell clumps that would affect the measurement, identified by visual inspection of all plates; 2) outliers among replicates; 3) epigenetic regulators that are lowly expressed in PDX MEXF 2090 (Log_{10} pseudocounts < 1) or PDX LFXL 1674 (Log_{10} TPM < 0.5) cells; 4) KO populations with severely compromised fitness in the unperturbed condition (20% reduction compared to the plate median for PDX MEXF 2090 and less than 20,000 cells at endpoint for PDX LFXL 1674). The last step was performed to avoid inflated stress/unperturbed ratios.

0.8% of the total imaged wells were discarded from the subsequent analysis and in more than 90% of the cases that was due to the presence of visible cellular clumps.

2.6.5 Data analysis

For each KO or control population the stress-specific fitness was derived from the stress/unperturbed ratio in cell count at endpoint. KO populations with enhanced or reduced fitness were defined based on the formula: $Z=(\chi-\mu)/\sigma$, where χ is the fitness of individual KO populations, μ is the mean fitness of negative controls, σ is the standard deviation of the fitness of negative controls. Populations exhibiting enhanced or reduced fitness were defined as those with a z-score > 1.645 or <-1.645 (90% confidence interval), respectively. Validation fitness assays were performed in a similar way, either measuring cell count over time using replicates fixed at various time points, or at endpoint. Population growth was quantified by normalizing the average cell count at each time point to the pre-treatment count (d0). In experiments where the population growth was monitored over several weeks, smoothing was applied to the curves to account for technical noise introduced by media change.

2.7 Treatment with epigenetic inhibitors

To identify drug concentrations 2-fold and 10-fold titrations were performed ranging from 20 μM down to 1 nM. A list of all the drug inhibitors used in this study along with their defined working concentrations can be found in Table 6. In the respective fitness assays and the FRET sensor live-cell imaging experiment [Figures 14, 32, 41, 42 and 43], cells were pre-treated for

3 days before being exposed to stress and drugs and media were replaced every 3 days to maintain efficient target inhibition over the course of the experiment. When assessing the reversibility of the stress-resistant phenotype [Figure 32], cells were grown for 9 days in the presence of the inhibitors and for an additional 9 days in the absence of the compounds.

2.8 Quantification of proliferative and apoptotic fractions

Two hours before the endpoint, populations were pulsed with 10 µM EdU and 5 µM of Ac-DEVD-NucView488 (Cen et al., 2008). Live cell imaging and quantification of Caspase 3 activity was performed using an Incucyte® S3 Live-Cell Analysis System. Subsequently, cells were fixed in 4% PFA stained for EdU by using the Click-iT™ EdU Cell Proliferation Kit (ThermoFisher Scientific, #C10340) as per manufacturer's protocol. Nuclei were labeled by adding SYTOX™ Green Nucleic Acid Stain (ThermoFisher Scientific, #S7020). Imaging and quantification were performed in Incucyte® S3.

2.9 *In vitro* clonogenic assays

For qualitative assessment, 1500 cells were plated in 6-well plates, while for quantification 10 cells/well were seeded in 96-well plates. After 24 hours, the medium was refreshed to RPMI 1640 without glutamine and cells were grown for 12 days until visible colonies appeared. Media was refreshed every 3 days. To achieve colony quantification, cells were fixed with 4% PFA, permeabilized with 0.5% Triton X-100 in PBS, stained with SYTOX™ Green Nucleic Acid Stain (ThermoFisher Scientific, #S7020) and imaged using an Incucyte® S3 Live-Cell Analysis System.

2.10 *In vitro* competition assays

2.10.1 Generation of MEXF 2090 cell lines stably expressing fluorescent proteins

MEXF 2090 cells were transduced with lentiviral constructs expressing GFP-NLS (pTRIP-SFFV-EGFP-NLS), or mCherry-NLS (pTRIP-SFFV-mCherry-NLS) (see also section 2.3.1). Seven days after infection, cells expressing similar levels of the fluorescent proteins were isolated by flow cytometry (see also section 2.12). mCherry-labelled lines were transduced with sgRNA constructs targeting the indicated epigenetic regulators, while GFP-labelled cells with sgRNAs targeting the non-expressed gene *TNP2*. KO populations were generated as described in section 2.5.

2.10.2 Co-culture of KO populations in nutrient deprivation

For the co-culture experiment, mCherry-labelled KO populations were mixed with GFP-labelled control cells in equal quantities and seeded in 96-well plates at a final density of 3000 cells per well. A mix of GFP- and mCherry labelled control cells was used as a baseline to account for possible differential fitness of the two labelled lines. Twenty-four hours after plating, the medium was refreshed with phenol-free RPMI lacking L-glutamine (ThermoFisher Scientific, #32404014). Live-cell imaging was performed at 12-hour intervals using an Incucyte® S3 Live-Cell Analysis System. During the experiment, the media was refreshed every 3 days. After 12 days under nutrient starvation, the mCherry to GFP ratio was calculated and normalized to the one before treatment.

2.11 Immunofluorescence microscopy

Note: Embedding of tumors in paraffin, subsequent sectioning and H&E staining was performed by the Experimental Histopathology Facility at the Francis Crick Institute.

Immunostaining of cultured cells was performed using standard protocols as described in (Monserrat et al., 2021) using anti-Lamin A/C (Santa Cruz Biotechnology, #sc-7292, 1:200), anti-H3K27me3 (EMD Millipore, #07-449, 1:400), anti-H4K16ac (Cell Signalling Technology, #2591S, 1:500) primary antibodies and relevant fluorescent secondary antibodies. Imaging was performed using either an IncuCyte® S3 system or an Axiovert Zeiss confocal microscope. For analysis of tumors, portions of tumors harvested from mice were fixed in 10% formalin and embedded in paraffin. Sections were deparaffinized with xylene and rehydrated in an ethanol gradient. Antigen-retrieval was performed for 20 minutes at 95 °C in citrate buffer. Slides were then blocked, incubated overnight with anti-H3K27me3 antibody (1:200) or phospho-S6 Ribosomal Protein (Ser235/236) (Cell Signaling Technology, #2211) at 4 °C, washed, incubated with the secondary antibodies (ThermoFisher Scientific, anti-rabbit Alexa flour 488, #A-21206 or anti-rabbit Alexa flour 568, #A10042) for 1 hour at room temperature, washed 3 times with PBS, incubated with DAPI and mounted with ProLong Gold Antifade Mountant (ThermoFisher, #P36934). Slides stained with only the secondary antibody were used as a negative control. Stained slides were imaged using an Olympus VS120 Slide Scanner and images were processed with QuPath-0.2.2 and quantified image J 1.45s.

2.12 Flow cytometry cell sorting (FACS)

For FACS sorting of melanoma cells to generate a) PDX MEXF 2090 GFP-NLS b) PDX MEXF 2090 MCHERRY-NLS and c) PDX MEXF GLUCOSE FRET SENSOR the following protocol was applied. Initially cells from a full confluent 15cm plate were trypsinized, counted and pelleted via centrifugation at 300g for 3min.The cell pellet was then resuspended in sorting media to a density of 5 million cells / mL. Sorting media consists of 15 mM HEPES buffer, 1% BSA, 2 mM EDTA, 100 U/mL DNase, 100 U/mL penicillin and 100 µg/mL streptomycin in PBS. Cell suspension was filtered through a 40 µM nylon mesh (Corning, #352340) to discard cell clumps and ensure that cells are in a single cell suspension. FACS sorting was performed using a FACSAria II (BD Biosciences) or S3e Avalon (Propel labs) flow cytometer and cells were isolated in 5 mL polypropylene tubes (Corning, #352063) containing recovery media (50% sorting media:50% FBS). All sorted populations were pelleted via centrifugation at 300 g for 5 min and plated in fresh RPMI media for a minimum of 24h prior to any experimental use.

2.13 *In vivo* competition assays

Note: Intradermal injections of cancer cells were performed with the help of Cristina Morales Torres (Scaffidi lab). Subsequent injections and handling of the mice were carried out by the Biological Research Facility at the Francis Crick Institute. Next-generation sequencing was performed by the Advanced Sequencing Facility and processing of the generated raw data by Harshil Patel. Subsequent analysis was performed by Ioannis Loukas. Part of this methods section was written by Harshil Patel (Loukas et al., 2023).

MEXF 2090 cells transduced with sgRNAs targeting either *EZH2* or *TNP2* (non-expressed control gene) and treated with doxycycline for 10 days were mixed at equal ratio and injected in NSG mice obtained from the common Francis Crick colony. 5 x 10^5 cells from the mix were intradermally injected in both flanks of 11-13 week-old male NSG mice in 50 µl of PBS. Approximately 3 weeks after injection mice were randomly segregated in two groups and treated with either Bevacizumab (Stratech, #A2006-SEL-5mg; twice a week i.p. at the dose of 2 mg/kg or 8 mg/kg) or vehicle. Tumor volume was measured twice a week using electronic calipers until animals were humanely killed approximately 5 weeks after the first injection. At the endpoint, tumors were harvested and the relative abundance of *EZH2*-KO and *TNP2*-KO cells was estimated by next-generation sequencing of the sgRNAs amplified from tumors.

Before lysing the tumors for the subsequent analysis some representative slices where isolated for immunocytochemistry analysis (see section 2.11). The remaining parts of the tumors were then cut into small pieces of ~2–3 mm in length, transferred into gentleMACS M tubes (GentleMACS, #130-096-335) containing two volumes of ATL buffer supplemented with 1:10 proteinase K (Qiagen, #69506) and blended at high speed with a gentleMACS dissociator (RNA_01.01 Program). Subsequently, gDNA from tumors and the injected cells was extracted using a DNeasy Blood & Tissue kit as per manufacturer's protocol (Qiagen, #69506). NGS libraries were prepared by performing a two-step nested PCR using Herculase II Fusion Enzyme Kit (Agilent Technologies, #600679). The primers used are listed in Table 5. To ensure efficient amplification of the sgRNAs, multiple PCR reactions were run for each sample, using a maximum of 1 µg gDNA in 50-µl reactions with 20 cycles of amplification. Following the first round, the PCR product was cleaned using QIAquick PCR Purification Kit (Qiagen, #28104) and 1/50 of the reaction was used as template for the second PCR, run for 28 cycles. Final products were run on a 2% agarose gel and purified using a QIAquick gel extraction kit (Qiagen, #28706).

Libraries were sequenced on the Illumina MiSeq using the MiSeq Reagent Nano Kit V2 (Illumina, #MS-102-2001) with 250 bp paired end reads and generated approximately 6000 251bp reads per sample. Raw reads were trimmed with the fastx_trimmer tool available within the FASTX-Toolkit (version 0.0.14) http://hannonlab.cshl.edu/fastx_toolkit using the parameters "-f 122 -l 141 -m 20" to extract the sgRNA sequence. These were then mapped to a reference consisting of the 14 guide sequences of interest using BWA (version 0.5.9-r16) with the parameters "-l 20 -k 4 -n 4". sgRNA counts were obtained after filtering the mapped reads for those that had zero mismatches and mapped to the sense strand of the guide sequence. To quantify the relative abundance of sgRNAs in each condition, raw reads for each sgRNA were normalized to the overall read counts. It is worth mentioning that similar sgRNA ratios were observed after allowing for 0, 1, 2, or 3 mismatches (data not shown).

2.14 Live-cell imaging of metabolic states

2.14.1 Generation of MEXF 2090 FRET-sensor line

MEXF 2090 were modified to stably express the FRET-based glucose biosensor, which signals the presence of intracellular levels of glucose and has previously been reported as a solid indicator of the relative preference between glycolysis and OXPHOS in living cells (Kondo et al., 2021). The PiggyBac transposon containing the FRET biosensor was co-transfected with a plasmid expressing the PB transposase (PBase) (Liang et al., 2009) at a 1:3 ratio using the FuGENE® HD reagent as per manufacturer's protocol (Promega, #E2311).

The levels of cintrine expression were monitored over the next days in Incucyte® S3 Live-Cell Analysis System. A significant drop was detected after 7 days due to the fading of transient overexpression. To enrich for stable integration events, citrine-expressing cells were isolated by flow cytometry (for a detailed description of FACS sorting refer to section 2.12).

2.14.2 Imaging and quantification

Note: Image acquisition was performed in collaboration with Colin D.H. Ratcliffe (Sahai Lab, The Francis Crick Institute). Subsequent image analysis and quantification was performed by Ioannis Loukas. Part of this methods section was written by Colin D.H. Ratcliffe (Loukas et al., 2023).

For live-cell imaging, cells were plated in optical grade 96-well microplates (GBO, #655090). Two hours before imaging, the media was replaced with phenol-free RPMI with or without L-glutamine (ThermoFisher Scientific, #11835063 and #32404014, respectively). Cells were imaged at 15-minute intervals for 24 hours using an inverted Zeiss LSM 880 confocal microscope and Zeiss Zen software (v2.3). A Plan-Apochromat 20x/0.8 NA objective lens was used and the emission signals were detected using the internal 32-channel GaAsP detector. Excitation light from an argon ion laser set to 3.5% was passed through a 458/514/561/633 multiple beam splitter and emission light was detected between 464-506 nm for eCFP and 517-571 nm for sensitized emission (FRET). Image acquisition settings were set to 512 x 512 pixels, zoom 0.6, 8.24 µs pixel dwell and line 4 averaging. Master gain for eCFP detection was set to 850 and 750 for citrine detection. Digital gain was set to 1.0 for both and digital offset was set to 0 for both. 3 x 2 tiling with 5% overlap followed by stitching was used to capture a rectangular field of view. The FRET ratio per cell was calculated from perinuclear areas (4x4 pixels) by dividing the total intensity in the FRET channel by the total intensity in the eCFP channel. For time-lapse single cell analysis, individual cells were manually tracked over time and the FRET ratio at each time point was calculated as described above. The measurements from time points overlapping with active cell divisions were discarded and replaced with the average of 6 time points (three prior and three after) flanking mitosis.

To assess mitochondrial activity, cells were pulsed with 100 nM of Tetramethylrhodamine ethyl ester perchlorate (TMRE, ThermoFisher Scientific, #T669). After 30 minutes, TMRE fluorescence was imaged and quantified using an Incucyte® S3 Live-Cell Analysis system.

2.15 Next-generation sequencing (experimental set-up)

2.15.1 Exome-seq

Note: Exome sequencing was performed by the Advanced Sequencing Facility at the Francis Crick Institute. This methods section was written by Robert Goldstone (Loukas et al., 2023).

For whole exome sequencing, gDNA samples were quantified using a QuantiFluor dsDNA system (Promega, #E2670) on the GloMax Multi Detection System following the manufacturer's guidelines. Sample quality was assessed using the Genomic DNA ScreenTape System run on the TapeStation 4200 according to the manufacturer's instructions (Agilent Technologies, #5067-5365; #5067-5366). Subsequently, gDNA was fragmented using Covaris LE220-plus focused ultrasonicator and then prepared into libraries using the NEBNext Ultra II DNA library prep kit (NEB, #E7645S) according to manufacturer's instructions. Libraries were then combined in 8-plex and exonic regions enriched via hybridization using the Twist Human Core Exome kit according to the manufacturer's instructions (#100578; #101174; #100254). Libraries were sequenced on the HiSeq 4000 using paired end 100 bp reads.

2.15.2 Bulk RNA-seq

Note: Bulk RNA sequencing was performed by the Advanced Sequencing Facility at the Francis Crick Institute. This methods section was written by Robert Goldstone (Loukas et al., 2023).

Bulk RNA sequencing was performed in biological triplicates of 6 different KO populations of MEXF 2090 cells in which *EED, EZH2, HIST1H1B, SMARCD1, SUZ12* and the non-expressed gene *TNP2* were targeted and that were grown under nutrient deprivation for 12 days. Four and six replicates of unperturbed and nutrient-deprived cells, respectively, grown in distinct wells of 96-well plates were pooled and total RNA extraction was performed using a RNeasy Plus Micro kit (#74034). Total RNA was quantified using an RNA QuantiFluor RNA system (Promega, #E3310) on the GloMax Multi Detection system following the manufacturer's guidelines. RNA quality was assessed via the High Sensitivity RNA ScreenTape using the TapeStation 4200 (Agilent Technologies, #5067-5579). RNA was normalized to 30 ng and used for cDNA synthesis and library preparation using the QuantSeq 3' mRNA-Seq FWD kit

(Lexogen, #015) according to the manufacturer's instructions. Libraries were sequenced on the Illumina HiSeq 4000 with single ended reads.

2.15.3 scRNA-seq

Note: The single cell RNA sequencing experiment was performed by the Advanced Sequencing Facility at the Francis Crick Institute.

MEXF 2090 cells were transduced with lentiviral constructs expressing sgRNAs whose scaffold was modified to contain the Capture Sequence 1 that is compatible with the 10x Genomics Feature barcoding technology (Table 5). To maximize editing efficiency, the KO populations also contained the respective pools of unmodified sgRNAs sourced from the available sgRNA-library (Henser-Brownhill, Monserrat & Scaffidi, 2017). Six replicates for each KO population, grown in different wells of 96-well plates, were cultured under nutrient deprivation for 12 days and pooled at end point. A reverse time-course scheme was followed to allow simultaneous collection of d0 (unperturbed) and d1, d2 and d12 samples (nutrient-deprived) (Supplementary Fig. 7a). Control and KO populations from each timepoint were pooled together prior to the library preparation step to minimize technical variability. Single cell mRNA-Seq was carried out using the 10x Single Cell Gene Expression kit v3.1 with Feature Barcoding technology for CRISPR Screening according to the manufacturer's instructions (10x Genomics, #CG000205). Briefly, cell suspension was counted and assessed for viability using an EVE automated cell counter (NanoEntek). Approximately 10,000 cells (mix of control and KO populations), per time point, were loaded into the 10x Chromium chip. GEM generation, barcoding, cDNA synthesis and clean-up was carried out as per the 10X protocol. Subsequently, the gene expression and feature barcoding libraries were separated by size selection and sequenced on the HiSeq 4000 according to the 10x guidelines. For the gene expression and the CRISPR libraries, approximately 336 and 268 million reads were acquired, respectively (mean values per time point). This resulted in a sequencing depth of ~87.000 reads/cell and subsequent detection of ~ 5800 genes/cell. Multiple sequencing runs were performed and aggregated using the built-in pipeline of 10x Genomics' software (CellRanger 3.0.2).

2.16 Computational analysis

2.16.1 Exome-seq analysis

Note: Processing of the raw data generated by the Exome-seq experiment was performed by Phil East. Subsequent analysis was performed by Ioannis Loukas. This methods section was written by Phil East (Loukas et al., 2023).

Trimmed (Trim Galore v0.6.4_dev) 101 basepair paired-end reads were aligned to human genome reference sequence GRCh38 using BWA mem (v0.7.17-r1188) (Vasimuddin et al., 2019). A mean alignment rate of 68,686,981 properly paired read pairs per sample was obtained across the 10 samples (SD 17,806,612) with a mean insert size of 175.14 (mean SD 69.13). The mean coverage across at least 50% of the exome was 176.8. Duplicates were marked using GATK MarkDuplicatesSpark (v4.1.7.0) (Van Der Auwera et al., 2013) and base quality scores recalibrated using GATK BaseRecalibrator and ApplyBQSR (GATK v4.1.7.0). SNVs and indels were called using Strelka2 (v2.9.10) (Saunders et al., 2012). Genome sequence, indices and dbSNP calls were obtained from the GRCh38 GATK bundle. SNVs and Indels were annotated using SnpEff (v4.3t, SnpEff DB version GRCh38.86) (Cingolani et al., 2012). The nfcore/sarek pipeline (v2.6.1) (Ewels et al., 2020; Garcia et al., 2020) with Nextflow (v20.11.0-edge) (Di Tommaso et al., 2017) was used to run the analysis end to end. Read depth varied from 3 to 671 per gene. Mutations with a read count < 50 reads (lower 10%) for either allele were excluded due to low coverage.

2.16.2 Bulk RNA-seq analysis

Note: Processing of the raw data generated by the bulk RNA-seq experiment was performed by Harshil Patel. Subsequent analysis was performed by Ioannis Loukas. This methods section was written by Harshil Patel (Loukas et al., 2023).

RNA sequencing was carried out on the Illumina HiSeq 4000 platform in multiple runs and typically generated ~11 million 76/101bp strand-specific single-end reads per sample. Adapter trimming was performed with cutadapt (version 1.9.1) (Martin, 2011) with parameters "--minimum-length=25 --quality-cutoff=20 -a AGATCGGAAGAGC". The RSEM package (version 1.2.31) (Li & Dewey, 2011) in conjunction with the STAR alignment algorithm (version 2.5.2a) (Dobin et al., 2013) was used for the mapping and subsequent gene-level counting of the sequenced reads with respect to hg19 RefSeq genes downloaded from the UCSC Table

Browser (Karolchik, 2004) on 7th June 2017. The parameters used were "--star-output-genome-bam --forward-prob 1". Differential expression analysis was performed with the DESeq2 package (version 1.12.3) (Love, Huber & Anders, 2014) within the R programming environment (version 3.3.1). An adjusted p-value lower than 0.01 was used as the significance threshold for the identification of differentially expressed genes. Pseudocount values estimated with the DESeq2 package were used to identify non- and lowly-expressed genes in MEXF 2090 cells, as transcript per million (TPM) values normalized to gene-length are not suitable for 3'-mRNAseq datasets. Expression values (TPM) for LXFL 1674 cells were sourced from the publicly available Oncotest-Charles River dataset.

2.16.3 scRNA-seq analysis

Note: Paolo Inglese performed a) pre-processing of the raw data, b) cell clustering and identification of KO-enriched or control-enriched subpopulations and c) run the txburst to estimate the bursting properties. Phil East applied the SCDE/PAGODA algorithm. The subsequent analysis was performed by Ioannis Loukas and Paola Scaffidi.

The detailed methods regarding the single-cell RNA seq experiment can be found in Loukas et al (2023). In the following subchapter the key steps of the analysis are summarized.

2.16.3.1 Pre-processing

a. Assigning cells to KO populations:

As stated earlier, the single-cell transcriptomics experiment was multiplexed, meaning that all the profiled KO populations were mixed before sequencing. Naturally, the first step during data processing was to remove cells where sgRNAs were not detected, as these cells could not be allocated to specific KO populations. For the remaining cells, the normalized sgRNA counts were used to fit a logistic regression model and cells with a predicted probability greater than 95% were assigned to the corresponding label. At this stage, 2632 cells with ambiguous cell identities were discarded (out of the total 15722 profiled cells).

b. Filtering:

To minimise the impact of technical noise on the subsequent analysis and interpretation of the results, the dataset was filtered for a) genes detected in less than 20 cells across all time points b) cells with overall small number of detected genes and c) cells with a high percentage of mitochondrial genes counts, an established indicator of low-quality cells in single-cell transcriptomics experiments. 467 cells were removed from the dataset at this stage.

c. Normalization:

The last step of the processing involved scaling normalization of the data to remove bias generated by technical differences in cDNA capture or PCR amplification per cell. Based on the assumption that such artefacts should affect all detected genes equally, a scaling factor was calculated per cell and subsequently used to normalize gene counts. The process affected the dataset minimally, confirming that the experimental design minimized the technical variance.

2.16.3.2 Differential gene expression analysis

To identify differentially expressed genes among samples of interest, cell-cycle corrected data were processed through relevant Seurat pipelines. Pairwise comparisons were performed across all dimensions of the dataset (KO vs Control populations at each time point and each population across time points). Statistical significance of the observed differences in gene expression was assessed by non-parametric Wilcox test followed by correction for multiple testing. FDR < 0.01 was used as the threshold to classify genes as differentially expressed.

2.16.3.3 Cell clustering and identification of KO-enriched or control-enriched subpopulations

To dissect population heterogeneity at the latest time point (day 12; d12) the following steps were performed [Figure 51]:

i) Pairwise clustering: a KO and the control population at d12 were clustered based on the expression of the most variable genes (VGs) within the merged cell populations. Principal Component Analysis (PCA) dimensionality reduction was applied to the merged cell cycle corrected VG data. Clustering was performed via the nearest-neighbors method through the relevant functions available in Seurat.

ii) Enrichment analysis: Quantification of the relative enrichment of KO and control cells in each defined cluster was performed by hypergeometric test for each cell type. Only clusters significantly enriched for at least one cell population were used in the following steps of the analysis. Selecting clusters that are only enriched for either control or KO populations, was based on the assumption that these differentially enriched clusters should be the ones driving the phenotypic differences observed in the fitness assays.

iii) Extraction of cluster-specific gene sets: for each KO- or control-enriched clusters the DEGs were identified against all other cells of the merged dataset. Genes with FDR

< 0.05 were considered differentially expressed and used to characterize the transcriptional identity of each cluster.

iv) The above steps were repeated for all pairwise comparisons at d12 (all KO populations vs control cells)

v) Defining shared signatures across subpopulations: hierarchical clustering of all the above signatures to identify gene signatures shared by KO- or control-enriched subpopulations. Signatures consisting of less than 50 DEGs or that corresponded to cell clusters with a size smaller than 5% of the pooled set size were removed.

vi) Meta-signature extraction: After defining the optimal number of clusters, the gene sets characteristic of these meta-clusters were extracted. These gene sets were denoted as meta-signatures to distinguish them from the cell cluster-specific signatures defined earlier (step iii)

vii) GSEA: Genes defining the KO-enriched or control-enriched meta-signatures were finally analyzed by GSEA to identify affected pathways.

2.16.3.4 Pathway-score estimation

Fitness and stress signatures were retrieved from the respective Hallmark gene sets in the MSigDB database. Genes present in all time points within the dataset were used for the analysis. Pathway scores were estimated as the first principal component scores of the cell cycle corrected counts of the whole dataset, comprising all cell populations and time points.

2.16.3.5 Gene set enrichment analysis (GSEA)

Gene set enrichment analysis of pre-ranked gene lists of either whole transcriptomes or identified meta-signatures was performed, focusing on MSigDB (version 7.4) hallmark, curated and GO gene sets. All parameters were kept as default except for enrichment statistic, which was set to classic. To assess the enrichment of signatures in HVGs and HFGs, the “compute overlaps” function of MSigDB was used, focusing on the hallmark gene sets and setting a threshold for biological significance at p-value ≤ 10^{-10}.

2.16.3.6 Analysis of gene expression variance

In this thesis, expression variance of single genes was quantified by using coefficient of variation (CV^2) as a metric. Within the generated dataset from the single cell transcriptomics experiment, CV is strongly anticorrelated with the mean expression, biasing the detection of

variable genes towards the ones that are lowly expressed (noisy). To remove this genome-wide trend and correct for additional technical (e.g. gene length) and cell-cycle related bias that affect inference of variance, the SCDE/PAGODA algorithm was applied to the raw dataset (Fan et al., 2016). The generated matrix contained for each gene the mean expression, the normalized variance and the percentage of expression across cells. Lowly-expressed genes with low mean expression (Avmodes < 10) in a given sample were discarded to avoid unreliable variance values. Finally, to define highly-variable genes (HVGs) a normalized variance threshold of 1.21 was used (top 5% in unperturbed control cells).

2.16.3.7 Transcriptional burst analysis

Burst kinetics were estimated using the approach described in Larsson et al (2019). As an input the raw RNA counts from the scRNA-seq experiment were used. A two-state model of stochastic gene expression was used to model the expression distribution at each gene and estimate the parameters (frequency and size) of the transcriptional bursts. In this model, each gene can fluctuate between two distinct states, that are either permissive or refractory towards transcription, *ON* and *OFF* respectively. Maximum likelihood (ML) inference was used to obtain estimates and confidence intervals on bursting parameters. The generated parameters are K_{on}, k_{off} and k_{syn}. k_{on} and k_{off} are indicative of the time (in units of mRNA degradation) of the ON and OFF state respectively, while k_{syn} represents the transcription rate when the gene is in the ON state. Burst frequency is represented by k_{on} and burst size is defined as the k_{syn}/k_{off} ratio.

While a two-state model likely does not capture all the variables affected by the complex biological perturbation induced by nutrient deprivation (Tunnacliffe & Chubb, 2020), I assumed that pairwise comparisons between cell populations are not affected by possible artifacts. It is important to state that the scRNA-seq dataset generated in this thesis does not have allele resolution. However, I assumed that any over- or under-estimation of the kinetics parameters would equally affect all compared samples. This is particularly relevant as the aim of this analysis was not to estimate absolute kinetic parameters, but to semi-quantitatively assess relative changes across conditions. In line with this assumption, the bursting parameters from the two distinct alleles profiled by Larsson et al. were highly correlated ($R^2 = 0.79$) [Figure 8], indicating that estimates from the combined alleles should not substantially affect comparisons across conditions (Larsson et al., 2019).

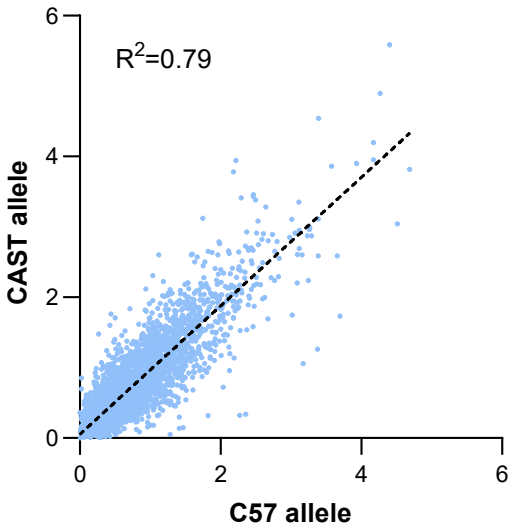


Figure 8: Allele specific correlation of burst frequencies

Correlation of burst frequencies as inferred via allele-specific single cell transcriptomics in primary mouse fibroblasts. Each dot represents a gene and its coordinates the burst frequency (k_{on}) as inferred in the CAST and C57 mice backgrounds. R-squared correlation from linear regression is shown. Raw data sourced from Larsson et al (2019).

The inference of the bursting properties was done for each sample (Cell type x time point) separately. The txburst algorithm returned accurate estimates of the bursting properties for at least 2000 genes in each sample (min: 2016, max: 3666). It is worth noting that these cohorts of genes were partially overlapping between samples (common genes in addition to genes where their estimate was possible only in one sample). For pairwise comparisons of the bursting parameters across the time trajectory only common genes sets were used. In the rest of the comparisons all the available accurately estimated genes were considered. High-frequency genes (HFGs) were defined in the control population (*TNP2*-KO cells) as genes with $K_{on} > 4$ (top 33%).

2.16.4 Statistical analysis

The type of statistical tests performed in this study, the value of N, and what N represents are indicated in the figure legends. Unless otherwise stated, all values are the average of individual values ± standard error of the mean (SEM) from at least three biological replicates. Statistical analysis was performed using either NGS-related packages or GraphPad software.

The choice among the different statistical test used throughout the thesis (e.g. T-test, ANOVA, Mann-Whitney, Kolmogorov-Smirnov, Wilcoxon signed-rank test, etc.) was determined per experiment based on the biological question and the nature of the dataset. The main

considerations that dictated the choice were i) the number of groups under examination, ii) if the groups are matched or unmatched iii) if the data follow a Gaussian distribution and iv) if the exploration within the samples focuses on detecting differences on the median or the overall distribution. In cases that I was exploring alterations in the mean behavior among two samples I used student's T-test. Characteristic examples are the difference in growth or apoptosis between cells cultured in unperturbed conditions or under nutrient starvation [Figures 11 and 14]. In other cases, I was interested in assessing alterations in the overall distribution of two unmatched groups and thus performed either Mann-Whitney or Kolmogorov-Smirnov (KS) tests. Examples of this, are the comparison of the FRET ratios between DMSO and EZH2i populations in response to stress [Figure 42] or the assessment of distribution changes in the expression of singles genes among different KO populations growing under nutrient starvation [Figures 53 and 63]. Finally, there were cases where I explored across multiple groups the overall effect between two variables (e.g. genotype and fitness under stress). In these cases, I used one-way ANOVA [Figures 22, 25, 26 and 58].

Chapter 3.
Selective advantage of epigenetically disrupted cancer cells under environmental stress

3.1 Introduction

Tumor progression can be considered an evolutionary process that is governed mainly by the following forces: a) diversity of biological properties b) selection that acts on diversity to promote the propagation of the fittest and c) genetic drift (Lipinski et al., 2016). Genetic and non-genetic alterations, along with their functional interplay, have been implicated in establishing intratumor heterogeneity in biological properties that ultimately fuels disease progression and relapse to therapy. Genome-wide sequencing studies of tumor specimens have detected recurrent loss-of-function mutations in epigenetic regulators. This disruption of the epigenetic network a) is a common phenomenon across different cancer types b) is affecting all functional classes of epigenetic regulation and c) is affecting multiple epigenetic regulators with varying recurrence (Brennan et al., 2013; Shen & Laird, 2013; Lawrence et al., 2014). On top of that, recent advances in our ability to infer when mutations are acquired during tumorigenesis demonstrated that epigenetic regulators are frequently inactivated in cancer subpopulations, uncovering an important role of epigenetic deregulation during subclonal expansion (Jamal-Hanjani et al., 2017; Dentro et al., 2021). Aside from potential gene-centric effects, the above patterns are suggestive of a model where disruption of multiple diverse components of the epigenetic machinery, at the network level, may converge towards similar cellular traits that are favorable and thus selected over time during the later stages of cancer evolution.

One important aspect of cancer evolution worth mentioning, is that cell fitness (what is considered favorable) is not a fixed property but is rather fluid and dependent on the specific context in which cancer cells grow, namely the tumor microenvironment (TME). This environment is characterized by the presence of hostile acellular properties (e.g. limited nutrients) and cellular components (e.g. immune cells) that are both spatially heterogeneous and dynamic over time (Junttila & de Sauvage, 2013; Wei et al., 2020). This spatial and temporal variability of unfavorable conditions creates the necessity for cancer cells to constantly survive and adapt to their changing environments in order to sustain uncontrolled proliferation that fuels tumor progression. Epigenetic regulators are among those regulatory layers that mediate the response of cells to external environmental stimuli, by dictating their transcriptional output. Thus, it is worth exploring whether their frequent inactivation can affect

the interactions of cancer cells with the tumor microenvironment and subsequently their ability to survive in this hostile milieu.

3.2 Aim

In this Chapter, I describe the steps taken to interrogate whether disruption of the epigenetic network alters the fitness of cancer cells under stress. Initially, I identify suitable cellular models that are devoid of mutations in epigenetic regulators and explore their response to various environmental challenges relevant to cancer cells. Next, I demonstrate in detail the optimization steps towards the successful completion of the large-scale fitness assays and the subsequent analysis of the generated data. Finally, I present the major observations, validate the identified gene-stress fitness relationships, and perform additional experiments to strengthen the emerging link between epigenetic deregulation and increased fitness of cancer cells under stress.

3.3 Results

3.3.1 Establishing experimental models of cancer cell survival to environmental challenges

3.3.1.1 Selecting distinct cancer models with minimal disruption in the epigenetic machinery

The first step towards assessing the effect of epigenetic deregulation in cancer cell survival under stress was to identify cancer models that are devoid of mutations in epigenetic regulators. The models used in this thesis are cell lines derived from Patient-Derived Xenografts (PDXs) from various cancer types. These populations more faithfully recapitulate genetic, histological and phenotypic features of the original tumors compared to the conventional cancer cell lines (Tentler et al., 2012). The available PDX-derived lines in our lab

are characterized by inactivating mutations in Tp53 and oncogenic driver mutations in KRAS/NRAS (KP cancer models, Table 3).

To explore fitness relationships and subsequent molecular mechanisms that are shared across cancer types, I selected two distinct cancer cell models: melanoma, of melanocytic origin and NSCLC lung carcinoma, of epithelial origin (MEXF 2090 and LXFL 1674 respectively). The above lines express 278 and 275 core epigenetic regulators respectively [Figure 9]. Their expression is significantly correlated between the two cell lines (R^2=0.54), with a small fraction being specific for either cancer type. Analysis of whole exome sequencing data obtained from MEXF 2090 and LXFL 1674 revealed that the epigenetic regulators investigated in this study are unaffected by loss-of-function mutations (gain of stop codon), making them suitable systems to dissect the functional consequences of disrupting the network through experimental gene inactivation.

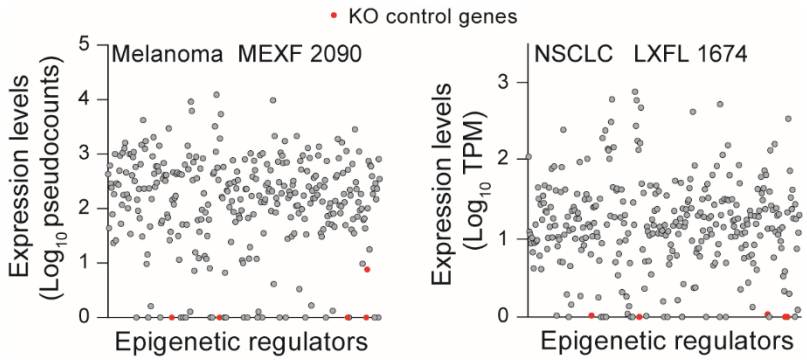


Figure 9: Transcriptomic characteristics of epigenetic regulators in MEXF 2090 and LXFL 1647 cells

Expression levels of epigenetic regulators in the indicated cancer models. Each dot is a gene. Genes selected as negative controls for CRISPR-induced editing are labelled in red. TPM: Transcripts per million.
KO control genes: non-expressed genes, targeted in the large-scale fitness assays as control

Disruption of specific genes will be achieved by inducing loss-of-function mutations in epigenetic regulators (ERs) through CRISPR-Cas9 gene targeting (see section 3.3.2.1). This creates the requirement for Cas9 expressing cancer populations. A previous member in the Scaffidi lab (Louise Richardson) generated clonal populations from the above PDX cell lines, inducibly expressing Cas9 nuclease upon treatment with Doxycycline. Gene knock-out (KO) is induced by transduction of sgRNA-expressing lentiviral constructs into Cas9-expressing cells. Despite the selection of transduced cells by antibiotic treatment, gene KO may not occur in all cells within the population. To estimate the KO efficiency, melanoma and lung cancer cells were transduced with an sgRNA targeting *LMNA*, a non-essential gene, whose protein

levels are easily quantifiable by immunofluorescence. 10 days after infection more than 65% of cells in both lines were negative for Lamin A/C, indicating efficient KO within the population and the existence of a clear window to detect phenotypic differences [Figure 10A]. To provide proof of principle of successful gene editing towards epigenetic regulation in the cancer models used, I next focused on Male Specific Lethal 3 (MSL3) which is a subunit of the MSL complex that catalyzes the acetylation of H4 at Lys-16. Quantifying the levels of H4K16Ac in *MSL3*-KO melanoma cells revealed similar KO efficiency to the one observed with Lamin A/C [Figure 10B]. It is worth noting that the presence of ~30% of cells escaping KO suggests an underestimation of the actual differences that will be detected between knock-out and control populations in the subsequent experiments.

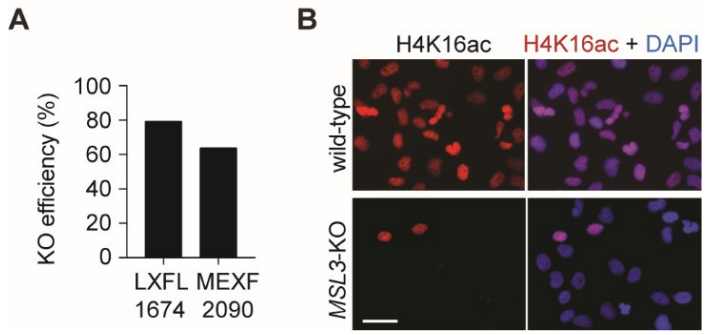


Figure 10: Robust gene inactivation in MEXF 2090 and LXFL 1647 cells

[A] Quantification of KO efficiency in the indicated cell lines transduced with *LMNA*-targeting sgRNAs as assessed by loss of Lamin A/C staining.
[B] Immunofluorescence microscopy detecting loss-of-function of the histone acetyltransferase MSL though loss of H4K16ac in polyclonal populations of MEXF 2090 cells transduced with *MSL3*-targeting sgRNAs. Scale bar: 10 μm

In summary, I selected two distinct cancer models (Melanoma and Large Cell Carcinoma) for the subsequent experiments. Their common driver mutations, clonal nature and the fact that they are devoid of major lesions in epigenetic regulators, make them suitable models to assess the effect of epigenetic deregulation.

3.3.1.2 Characterizing the response of cancer cell lines to nutrient starvation

Decreased cancer cell fitness under nutrient starvation

Aiming to probe how epigenetic deregulation affects cancer cell survival to unfavorable conditions, the next step was to incorporate into the experimental setup a stress condition relevant to cancer cells. During tumor progression, uncontrolled proliferation of cells is accompanied by suboptimal vascularization. This results in an imbalance between supply and

demand for resources, thus creating tumor species that face a substantial shortage in both oxygen and nutrients (Wei et al., 2020). One of the most limited nutrients within solid tumors is glutamine, a non-essential amino acid that plays a crucial role both in protein and nucleic acid biosynthesis (Kamphorst et al., 2015; Pan et al., 2016; Yoo et al., 2020).

To mimic the nutrient scarcity faced by proliferating cancer cells, I starved melanoma and lung cancer cells from L-glutamine and followed their behavior over time through time-lapse imaging. Both cell lines suffering glutamine deprivation exhibited a striking decrease in proliferation over 3 days [Figure 11A-B]. The observed growth inhibition was similar when glucose was depleted from the media of melanoma cells, suggesting that limiting levels of distinct nutrients can elicit similar cellular responses [Figure 11C]. This phenotypic similarity is in line with previous reports, demonstrating that diverse nutrient stressors can lead to highly similar and overlapping molecular effects (Gameiro & Struhl, 2018). To further characterize the behaviors underlying the decreased fitness of cancer cells during nutrient starvation[4] I employed a live-cell Caspase-3 fluorescent indicator along with EdU incorporation to directly assess the rates of apoptosis and proliferation respectively (Cen et al., 2008). Early response to nutrient starvation consisted of a gradual decrease in proliferation along with a respective increase in apoptosis. After 3 days under stress, these antithetic trends stabilized, reaching an equilibrium where approximately 20% of proliferating cells counteracted the death of a significant fraction of the population [Figure 11D], in accordance with the absence of population growth during the first week under starvation [Figure 12B]. Monitoring the cancer cells under chronic starvation (more than 3 weeks), revealed the presence of resistant colonies, likely due to the selection of favorable underlying traits, that could combat starvation and drive population growth under this unfavorable condition [Figure 11E].

Overall, nutrient deprivation poses a significant challenge for cancer cells. The initial reduction in proliferation is accompanied by a significant increase in apoptosis that plateaus over time. After prolonged culture, spontaneously stress-resistant colonies emerge able to reconstitute growing populations.

[4] Hereafter in this chapter, nutrient starvation refers to experimental procedures where cancer cells are deprived of the non-essential amino acid L-glutamine.

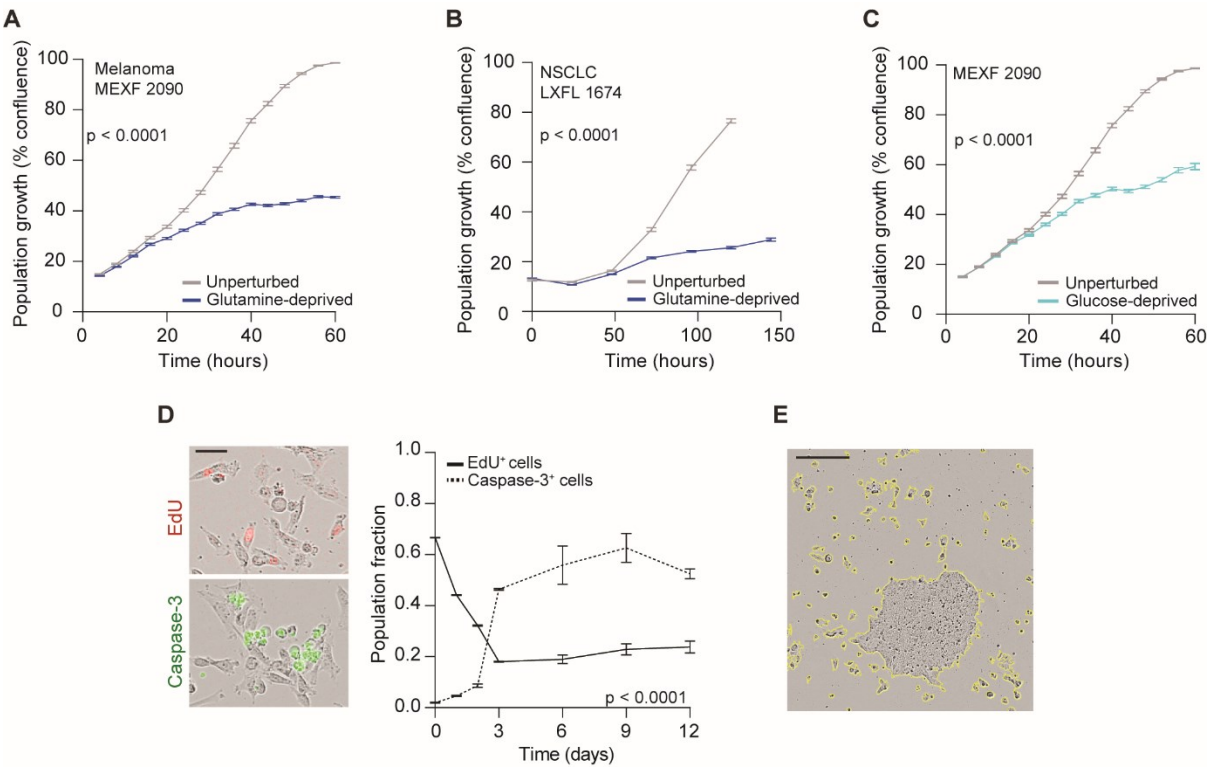


Figure 11: Nutrient starvation halts growth of cancer cells

[A] Growth kinetics of MEXF 2090 melanoma cells in unperturbed conditions or under glutamine deprivation. Values represent mean ± SEM from eighteen biological replicates. P-value from two tailed Student's t-test calculated at the last time point.
[B-C] Growth kinetics of the indicated cells in unperturbed conditions or under glutamine deprivation (B) or glucose-deprivation (C). Values represent mean ± SEM from twelve and eighteen biological replicates, respectively. P-value from two-tailed Student's t-test calculated at the last time point.
[D] Representative images and quantification of proliferating (EdU+) and apoptotic (Caspase-3+) MEXF 2090 cells grown under glutamine deprivation. Values represent mean ± SEM from three biological replicates. P-value comparing values at d0 and d12 for each curve from two-tailed Student's t-test. The fluorescent signal is overlaid on phase contrast. Scale bar: 50 μm.
[E] Phase contrast image showing expansion of a stress-resistant subpopulation of LXFL 1674 cells after 22 days of growth under glutamine deprivation. The outline of the IncuCyte cell mask is indicated by a yellow line. Scale bar: 400 μm

<u>Absence of mutational events mediating the phenotype of stress-resistant populations</u>

Note: Exome sequencing was performed by the Advanced Sequencing Facility at the Francis Crick Institute. Processing of the generated raw data was performed by Harshil Patel. Subsequent analysis was performed by Ioannis Loukas.

To explore in greater depth the capacity of cancer cells to survive under chronic starvation and interrogate the extent, probability and the time-frame of such behavior I generated 120 replicates, by seeding 3000 cells at the start of the treatment, and examined their response to

starvation over several weeks [Figure 12A]. While the initial response was homogeneous, at the latest time points (after 10 days under stress) the behavior varied significantly with some replicates displaying severely compromised growth, some showing early emergence of stress-resistant populations and others characterized by an intermediate phenotype [Figure 12B-C]. The difference in the mean response between the two cancer models can be attributed to the significantly lower overall proliferation rate that characterizes the lung cancer cells [Figure 12D]. Of note, a correlation was observed between the number of resistant colonies and overall population fitness, suggesting that the phenotypic differences observed within the replicates can be attributed to the selection of rare subclones [Figure 12E].

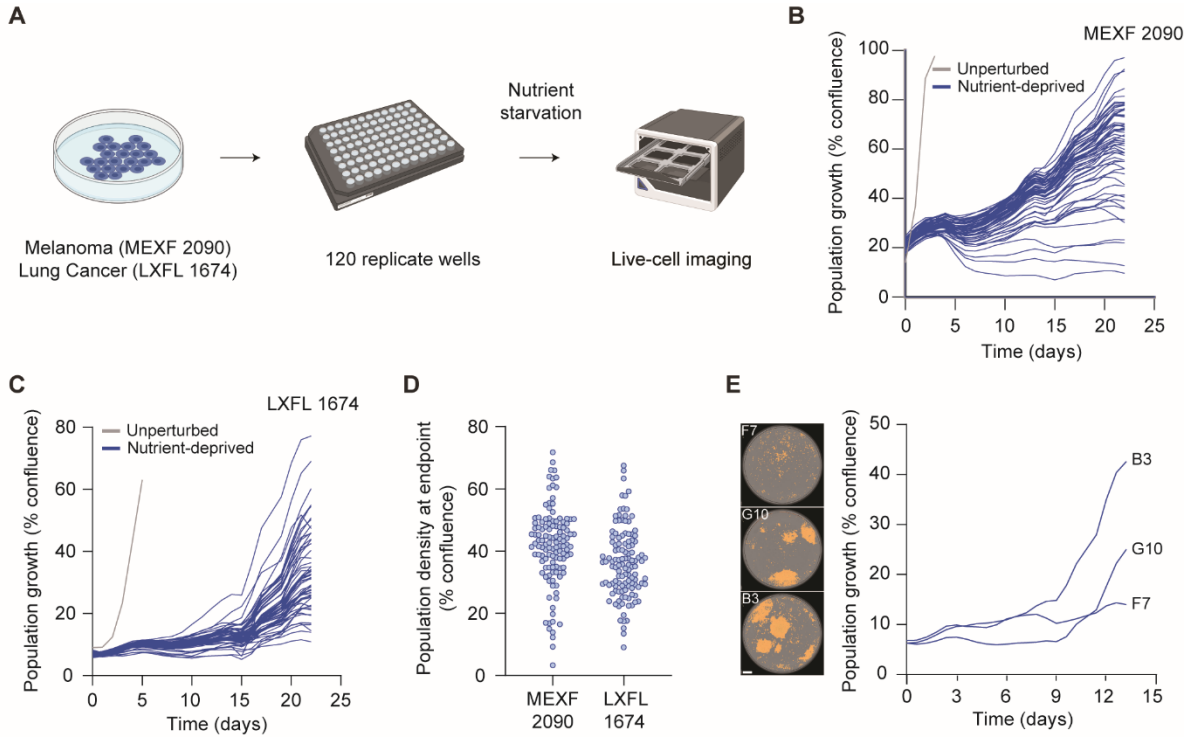


Figure 12: Variable response of cancer cells under prolonged nutrient starvation

[A] Schematic representation of the experimental procedure
[B-C] Growth kinetics of 60 replicates of LXFL 1674 (B) or MEXF 2090 (C) cells under the indicated conditions.
[D] Population density of individual replicates of the indicated cell lines after 22 days of growth under nutrient deprivation.
[E] Growth kinetics of three biological replicates of LXFL 1674 cells under nutrient deprivation, and corresponding images. IncuCyte mask capturing confluence within the wells (phase contrast imaging) is shown in yellow. Scale bar: 800 μm

The aforementioned heterogeneity can have either a genetic or non-genetic basis. The clonal nature of the lines used, the distinct behavior of the replicates and the stochastic emergence of stress-resistant subpopulations argue against a scenario where the selection of pre-existing

genetic events mediates the the observed phenotypes. To directly assess the potential contribution of mutations in the emergence of resistance to stress, I selected populations that managed to recover growth under nutrient deprivation and compared their mutational landscape via exome sequencing to the one before treatment [Figure 13A]. After accounting for technical noise, analysis of the mutational landscape detected few genetic alterations in the examined populations (min: 3, max: 18). However, none of these alterations were clonal. To the contrary, they were characterised by low frequency within the surviving populations and their abundance after exposure to stress correlated with the one in the treatment naïve parental populations [Figure 13B]. Further examination of these sub-clonal SNVs did not reveal any recurrent mutational events shared across different replicates of the same cancer type, limiting the ability to assign biological significance to these events [Figure 13C].

Although the above data cannot exclude the possibility that every single population acquires a different mutation (or combinations of them in different surviving subclones) that confers a survival advantage to starvation, they suggest that non-genetic events may be important for the observed phenotypes and are worth exploring.

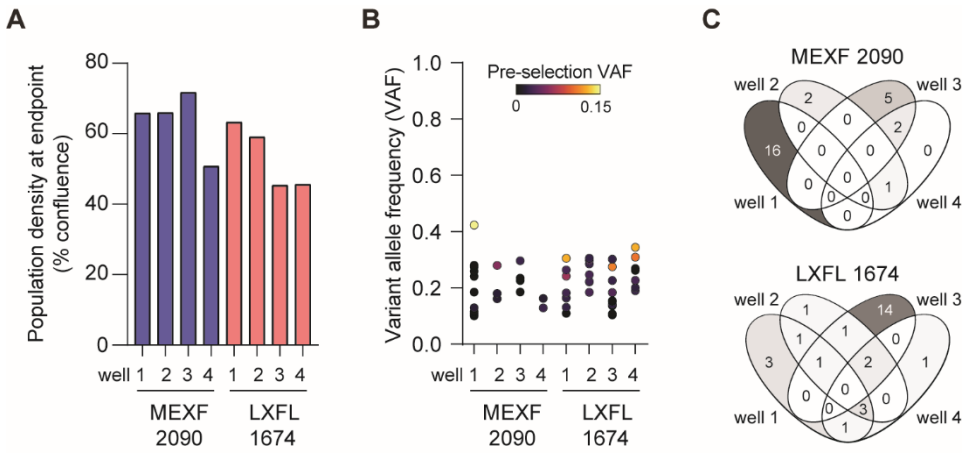


Figure 13: Absence of clonal or shared subclonal mutations in cancer populations surviving under nutrient starvation

[A] Population density of individual replicates of the indicated cell lines after 22 days of growth under nutrient deprivation.
[B] Clonality of the detected SNVs. The variant allele frequency (VAF) of each mutation in the parental population is indicated by colours, showing that detected mutations were already present at higher frequency before bottleneck selection. SNV: Single Nucleotide Variation
[C] Venn diagrams showing the number of subclonal mutations shared by the indicated stress-resistant subpopulations as indicated in B.

Inhibition of selected epigenetic regulators enhances cell survival under nutrient starvation

To get a first indication regarding the potential role of epigenetic regulation in the response to stress, I treated cancer cells with a panel of chemical compounds that inhibit the function of a variety of epigenetic regulators including histone deacetylases, methyltransferases and demethylases, with either repressive or activating functions. More specifically, melanoma cells (PDX MEXF 2090) were pre-treated for 3 days with the drugs. Considering that these cells are fast cycling (approximately doubling every 18 hours), this timeframe provided enough time for robust resetting of the epigenetic network (e.g. loss of histone marks upon inhibition of the respective enzymes, etc). Cells were then cultured either in the absence of stress or in starvation (L-glutamine deprivation) and their fitness was monitored over time. Multiple compounds increased cell fitness under nutrient deprivation, with treated populations showing up to 5-fold more cells than control populations after 9 days [Figure 14]. In the unperturbed conditions most of the treated populations exhibited comparable proliferation rates, with the exceptions of populations where HDACs were inhibited that demonstrated decreased growth [Figure 14]. This reduction in the proliferation rate was accompanied by an alteration in cellular morphology from a mesenchymal to an epithelial state. Overall, selective epigenetic inhibition can increase the fitness of cancer cells, and this advantageous effect is specific only under stress.

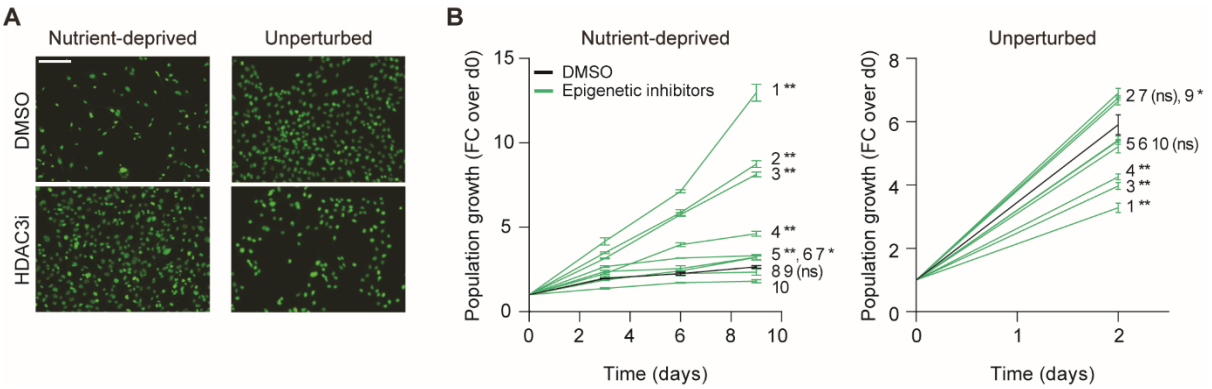


Figure 14: Chemical inhibition of epigenetic proteins enhances survival of MEXF 2090 cells under nutrient starvation

[A-B] Representative images of endpoint populations (A) and growth kinetics of MEXF 2090 cells treated with the indicated compounds and grown under the indicated conditions. Values represent mean ± SEM from three biological replicates. P-values from two-tailed Student's t-test (*p < 0.05 and **p < 0.01).
1: RGFP966, 2: GSK126, 3: Tubastatin A; 4: Quisinostat; 5: EPZ004777; 6: WM-8014; 7: EX527; 8: DMSO, 9: JIB-04; 10: MM102. Nuclei are visualized with SYTOX green nucleic acid stain. Scale bar: 50 μm

In summary, the data presented in this subchapter confirm that nutrient starvation poses a major challenge for cancer cells. Cancer populations devoid of inactivating mutations in epigenetic regulators exhibit significant phenotypic variability when grown under nutrient deprivation. While the majority of the cells suffer and die under stress, there is a small fraction that survives and drives long-term population growth under stress. Mutational analysis of these rare survivors failed to detect any genetic basis for their behavior. Finally, initial evidence suggests that interfering with epigenetic regulation can promote cancer cell survival under stress, hinting towards further investigation of this functional relationship.

3.3.2 Systematic disruption of the epigenetic regulatory network under nutrient starvation

Having selected the cancer cell lines and assessed their response to a characteristic property of the tumor microenvironment, nutrient scarcity, I then systematically disrupted the epigenetic network in both melanoma and lung cancer cells and looked for alterations in the capacity of the cells to survive under stress. The fitness of cancer cells is dependent both on their intrinsic proliferative capacity but also on their ability to withstand environmental challenges. A perturbation may lead to increased survival because it affects the intrinsic cycling capacity of the cells irrespective of the stress condition per se. In an alternative scenario, a mutation may exert an effect only under nutrient starvation, without affecting the proliferation of the cells in the absence of stress. To account for the above scenarios, I explored the behavior of the cells both in unperturbed and stressed conditions and quantified the stress-specific fitness. The scope of this approach was not to identify top hits, for example gene KOs that have the strongest effect on survival under stress, but rather explore patterns across the dataset to reveal potential network-level effects. This creates the necessity of detecting even mild phenotypes among the interrogated populations. Such cases could be masked in approaches involving pooled CRISPR screens, thus I opted for an experimental setup consisting of the inactivation of epigenetic regulators in an *arrayed* format. In the next subchapter, I am presenting the experimental procedure followed during the large-scale fitness assays along with the subsequent steps taken to ensure the quality of the generated dataset.

3.3.2.1 Large-scale fitness assay: design and experimental pipeline

Plate design

To induce loss-of-function mutations in the epigenetic machinery, I utilized an *arrayed* lentiviral CRISPR/Cas9 library that was previously developed in the lab (Henser-Brownhill, Monserrat & Scaffidi, 2017). This library targets 450 genes broadly involved in phenomena related to chromatin structure and function. On top of core epigenetic activities, the library contains genes that present sequence and functional domain similarity to epigenetic regulators along with genes involved in DNA Damage response and repair. Considering the scope of the current thesis, I selected 318 core epigenetic regulators to inactivate, belonging to 16 distinct functional classes [Figure 15A] (Table 1)

Pilot experiments in 96-well format, where the growth of control populations under nutrient starvation was examined, highlighted the need to include multiple control populations to account for technical variability in the large-scale experiments. Each library plate was therefore designed such that [Figure 15B]:

A. Outer wells are excluded to avoid the edge effect. Pilot experiments where replicates of identical populations were grown in unperturbed conditions, revealed that the phenotype obtained by populations in the outer wells of a 96-well plate was recurrently lower compared to the rest. This effect was magnified when replicate populations were cultured under stressful conditions (data not shown). Aiming to minimize such bias and increase the reliability of the detected phenotypes, I decided to exclude these wells from the final format.
B. Control populations occupy 1/3 of each 96-well plate. This substantial number of control wells allows for filtering out technical outliers during downstream analysis. The presence of non-targeting sgRNAs, commonly used as control samples in CRISPR screens, can exacerbate promiscuous editing thus significantly biasing the detected phenotypes. To minimize the introduction of such errors in the system, cells carrying sgRNAs against non-expressed genes were used as negative controls (Chen et al., 2018).
C. The selected negative controls carry sgRNAs against *TNP1, TNP2, HMGB4, SMC1B* or *DNMT3L*. Analysis of RNA-seq data revealed that the above genes are not expressed in either MEXF 2090 or LXFL 1647 cancer populations [Figure 9]. For each control population, four intra-plate replicates are assigned, to control for variability attributed to technical factors like disparities in initial plating, position effect within the plate and others. The inclusion of 5 distinct genes as controls buffers potential outlier effects in the phenotype of control populations (e.g. off target effects). The overall

number of negative controls per plate (20) provides a substantial amount of data regarding the mean behavior of control populations, accounts for potential outliers and thus allows for robust exploration of phenotypic deviations from the norm.

D. 1 well in each plate is occupied by *ARID2*-KO cells, which demonstrated significant fitness deviation under stress in the pilot experiments. The phenotype observed in these *ARID2*-KO populations across different plates is used as an indicator of reproducibility between different experiments.

Based on the above plate design, the PDX MEXF 2090 or LXFL 1647 KO library was distributed across eight distinct plates. After having refined the plate design and optimized step by step the high-throughput production of KO cells in 96-well format, I generated the libraries of KO populations for both melanoma and lung cancer models [Figure 15C]. Of note, I avoided prolonged culture of the newly generated KO populations (10 days from induction to fitness assays) to minimize the potential population diversification. The goal is to culture these plates in unperturbed or nutrient-starved conditions and ask if the epigenetically deregulated cells exhibit differential fitness under stress compared to the control populations [Figure 15D].

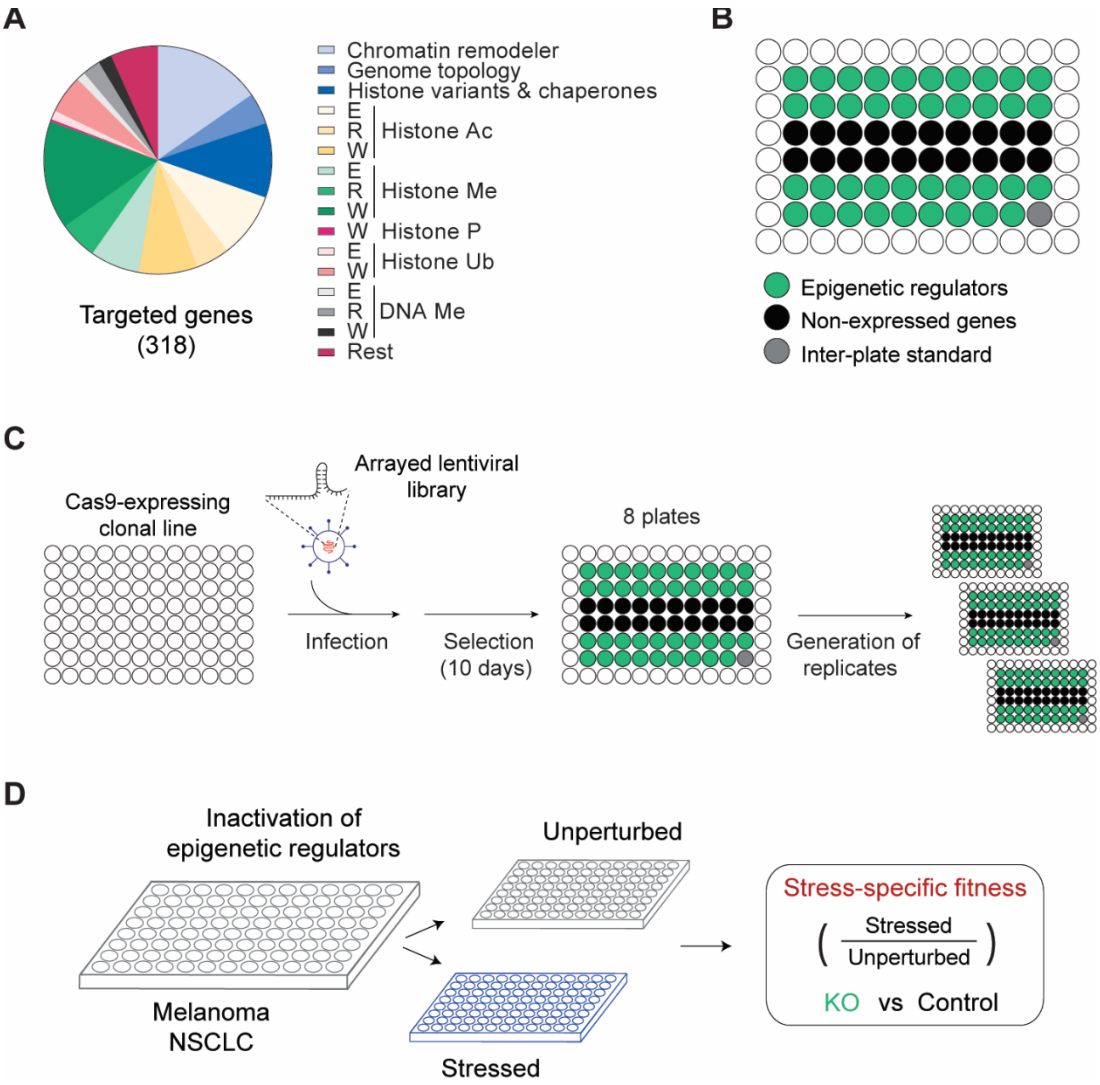


Figure 15: Systematic disruption of epigenetic regulators in MEXF 2090 or LXFL 1647 cells

[A] Representation of functional classes of epigenetic regulators among the genes targeted by the arrayed CRISPR library. E: eraser; R: reader; W: writer; Ac: acetylation; Me: methylation; P: phosphorylation; Ub: Ubiquitination.
[B] Library plate format. KO populations for epigenetic regulators occupy 66% of the plate while the rest is populated by control samples. Each plate carries an *ARID2*-KO population as an indicator of phenotypic reproducibility across plates.
[C] Schematic representation of the workflow followed for the generation of the melanoma or NSCLC KO library.
[D] Schematic of the experimental approach used to explore the effect of epigenetic deregulation on cancer cell fitness under stress.

Assay pipeline

The KO library (318 targeted genes) consists of eight plates. Based on the assay format, each library plate needs to be cultured in three different conditions (two different durations under nutrient deprivation and unperturbed). The scope behind selecting two distinct durations under stress is to be inclusive and capture both early responders but also detect survivors that deviate from the norm at later stages. To do so, during each experiment I generated multiple replicate plates from each library plate (> 15). Due to the scale of the experiments, during this seeding step it is not possible to count cells for each KO population. Thus, I opted to let the library plates reach confluence, allowing for different populations across the plate to reach similar densities. During splitting, a uniform dilution factor was applied across the plate aiming for approximately 10% density. This approach resulted in comparable plating density across the different populations within a library plate [Figure 16]. Approximately 24 hours post seeding, when the median cellular density of the plates reached the designated density of approximately 4000 cells, I exposed the cells to the defined stress conditions. Throughout the experiment, one representative plate was monitored by time-lapse imaging, to confirm that the growth kinetics was similar to the ones expected based on pilot experiments and also dictate the respective endpoint where the cells are quantified by nuclei staining. Finally, I processed the raw data to assess the fitness of each knock-out population in the different contexts relative to control populations, normalizing for potential stress-independent growth changes induced by gene KO (growth under stress/growth under unperturbed conditions) [Figure 16].

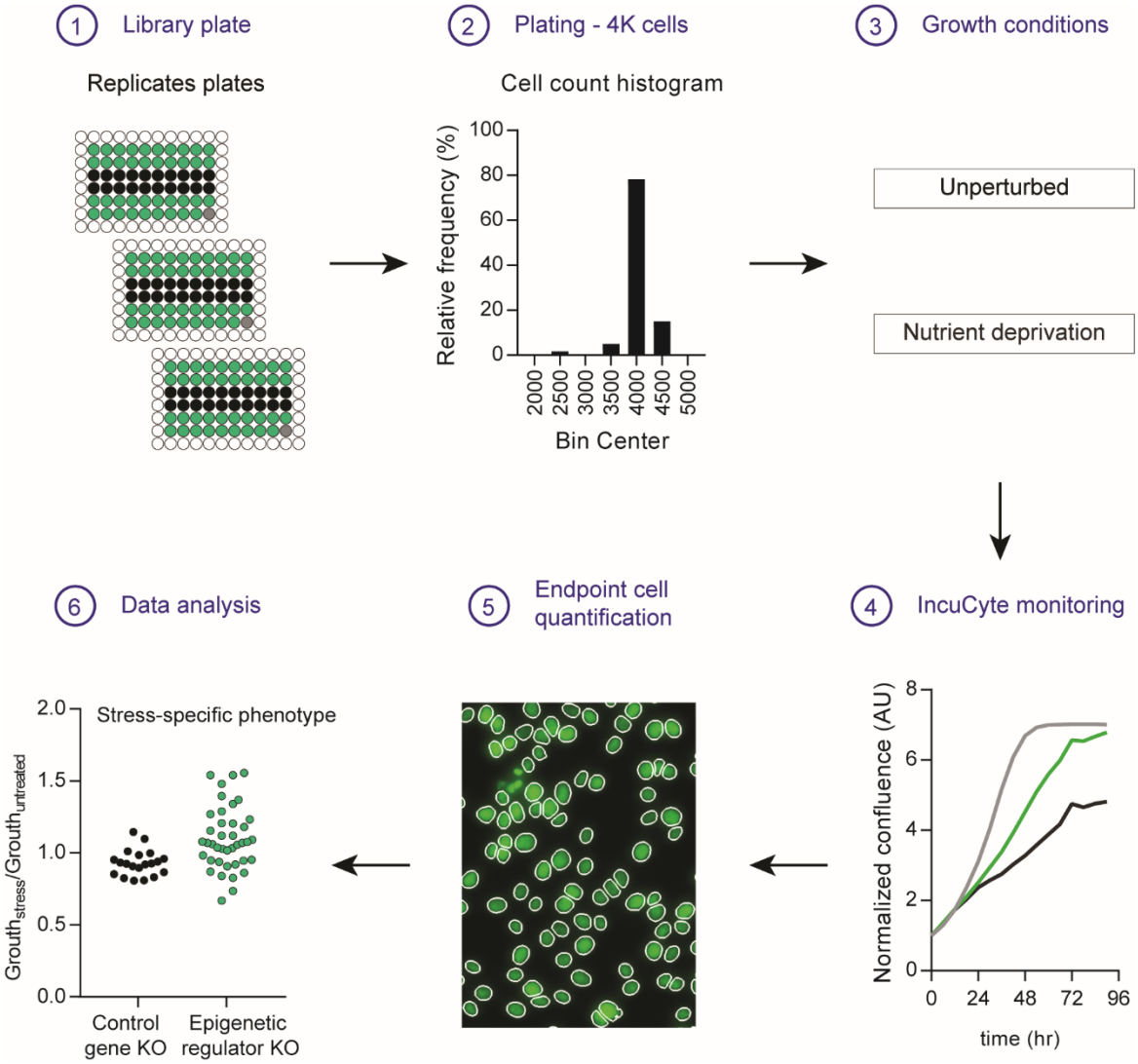


Figure 16: Experimental pipeline followed during the arrayed large-scale fitness assays

3.3.2.2 Inferring the stress-specific phenotype

The generated data from the large-scale fitness assays were then processed through a quality control pipeline, to ensure their validity before any further data analysis. Filtering of data included the following steps: a) removal of outlier replicate wells that are often characterized by cellular clumps that hinder accurate quantification of population phenotype b) removal of populations with severely reduced growth in unperturbed conditions, that would inflate the ratios of the inferred stress-specific fitness (see below) c) exclusion from the analysis of genes that are not expressed in the respective cancer models and thus their inactivation is not expected to result in phenotypic deviation. For greater details refer to Methods section 2.6.4. It this worth mentioning that the filtered data from step a represented less than 1% of the total imaged wells with the majority of the cases (>85%) being excluded due to the presence of

visible cell clumps. 262 and 250 populations of MEXF 2090 and LFXL 1674 respectively passed the filtering from steps b and c and were then used for downstream analysis.

Having confirmed the good quality of the raw data, I then went on to calculate the stress-specific phenotype for each KO population. To do so, in the first step of data processing, I divided the cell count of each population obtained under nutrient starvation over the mean of its untreated counterpart to determine the stress-specific fitness [Figure 17A-B]. This step corrects for technical variance irrespective of the stress condition. More specifically, all the experiments are characterized by technical noise as a result primarily of the variable pre-treatment density across different populations. This can be attributed to the way the seeding of the populations during the large-scale fitness assays was performed (see above). Importantly, nutrient starvation exerts its effect on cells in a density-independent manner**.** In pilot experiments where different numbers of parental melanoma cells were seeded and then exposed to nutrient starvation, normalization over the mean growth of unperturbed populations could correct for this discrepancy, allowing for solid comparison between samples (data not shown).

By definition, the control populations should demonstrate the same phenotype, thus observed differences among them indicate the presence of technical artifacts. By normalizing over the growth in the unperturbed condition, I correct for such technical discrepancies, evident by the narrow distribution of the negative controls after this step of data processing [Figure 17B]. It is worth mentioning that this approach was selected based on initial pilot experiments, where different normalization methods were assessed. It primarily corrects for differences that emerge from the variability in the initial plating density across the different populations within the plates. To avoid inflated ratios that could bias the analysis, populations within significant deviation from the norm in unperturbed conditions were excluded from the downstream analysis (See Methods Section 2.6.4 for more details). Next, I searched for populations that demonstrate stress-specific fitness above or below the scatter of the control samples, characterizing them as advantageous or disadvantageous respectively [Figure 17B]. To mathematically estimate if and how different is the phenotype of each KO population compared to the behavior of the controls I employed the z-score metric. Z-score describes how far a raw measurement is (number of standard deviations) from the mean of a group of measurements and acts on the assumption that the data are sampled from a normal distribution. Thus, based on the distribution built from the data of the control populations I defined a threshold (90% confidence interval) that could be used to classify the observed fitness deviations as neutral, positive, or negative. [Figure 17C]. To ease inspection of the generated dataset, I finally visualized the stress-specific fitness for each gone as a heatmap [Figure 17D].

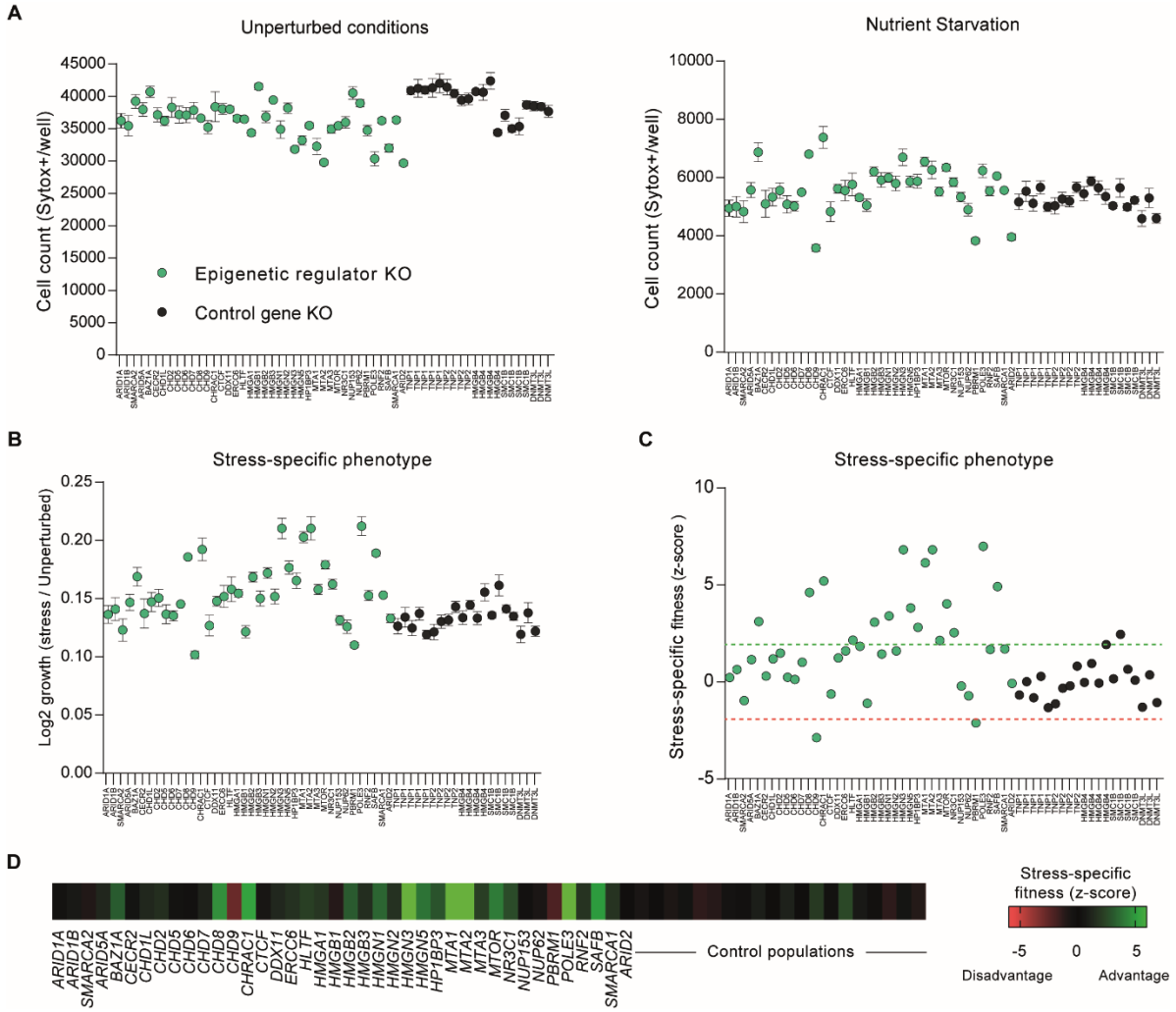


Figure 17: Inferring the stress-specific phenotype

[A] Dot plots of nuclei count in unperturbed (left) or nutrient starved conditions (right). Each dot indicates the nuclei count per KO population (Mean ± SEM).
[B] Normalization of the quantification obtained in a stress context over the unperturbed one infers the stress-specific phenotype.
[C] Stress-specific fitness of distinct KO populations under nutrient starvation. Green and red lines indicate the z-score thresholds used to define populations with enhanced or decreased fitness, respectively.
[D] Heatmap visualization of the stress-specific fitness.

3.3.2.3 Consistency and reproducibility of the dataset from the large-scale fitness assays

I conducted in total eight different large-scale fitness assays, one per library plate, following the above experimental and data analysis pipelines. Within each experiment, the technical reproducibility was satisfactory, evident by the negligible deviation between replicate measurements in the presence or absence of stress [Figure 18A]. This minimal variance

ensures high sensitivity of the assay allowing the detection of even subtle phenotypic deviations. Across different experiments, the phenotype exhibited by the control populations revealed a) overall consistency across plates, b) the presence of technical outliers in each plate that diverge from the norm, in line with initial observations in pilot experiments c) differential scatter of the controls across plates. These observations reinforce the importance of incorporating a substantial amount of control populations within each library plate, to robustly normalize for technical variability [Figure 18B].

Despite this intrinsic variability, multiple lines of evidence suggest that the detected phenotypes are biologically meaningful.

A. Inter-plate phenotypic reproducibility.

   ARID2-KO populations demonstrated a strong fitness disadvantage under nutrient starvation and an advantage in conditions where cells are challenged by induction of replication stress (an additional stressor that is presented in greater detail at section 3.3.3.5). Thus, it was selected as an inter-plate standard (present at the same position in each library plate), to assess phenotypic reproducibility. A similar pattern was observed in the majority of library plates indicating that the assay pipeline can robustly detect phenotypes in a reproducible way. It is important to note though that the extent of the detected fitness leap can vary, arguing that technical aspects can affect the degree of the detected phenotypic outcome [Figure 18C].

B. Members of the same protein complex demonstrate similar phenotypes.

   EZH2, SUZ12, and EED are components of the Polycomb repressive complex 2 (PRC2). The phenotype detected in populations KO for either of these subunits was highly similar [Figure 18D]. Comparable patterns were also observed in other sets of interacting proteins, like the members of NuRC and WCRF/CHRAC complexes. Although such relationships could emerge by chance when sampling from a pool of more than 300 different KO populations, they represent good indications of phenotypic reproducibility.

C. Neutral fitness upon inactivation of non-expressed genes.

   Genes that are not expressed in MEXF 2090 cells demonstrated phenotypes comparable to the control populations [Figure 18D]. In line with the above observation, genes that did not present any deviating phenotype from the norm in any stress condition (Neutral) have significantly lower expression levels compared to the ones that exhibited fitness advantage or disadvantage in at least one condition (Non-neutral) [Figure 18E]. The above observations are correlative in nature. Thus, it is important to

note that they represent indications and not proof of the robustness of the large-scale fitness assays.

Comparable results were also obtained from the large-scale assay performed in LXFL 1674 cells. Overall, this initial interrogation of the data, suggests that the large-scale assays in both melanoma and lung cancer cells represent a solid source of information worth further investigation.

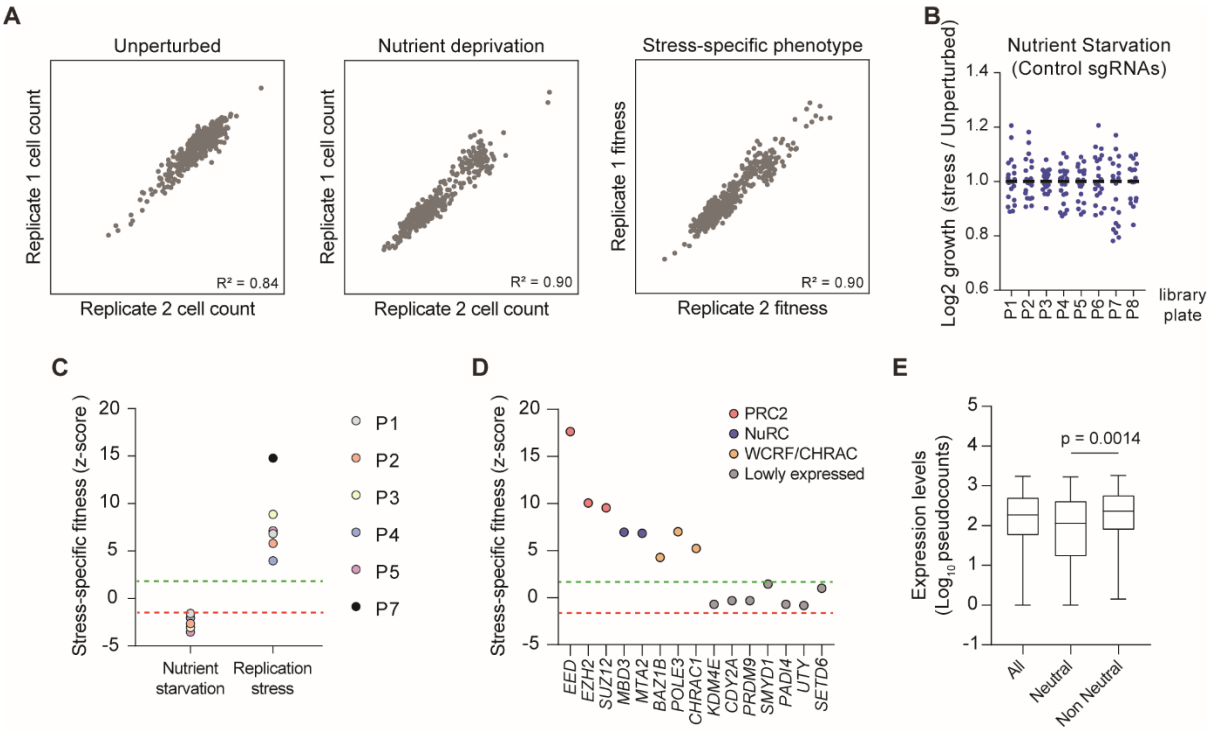


Figure 18: Reproducibility of the data generated by the large-scale fitness assays

[A] Correlation between biological replicates of MEXF 2090 cells in the indicated conditions. The coefficient of determination R2 is indicated.
[B] Phenotypic scatter of control populations across different library plates. Dots represent the mean stress-specific phenotype of a control sample.
[C] Quantification of stress-specific fitness for *ARID2*-KO replicates of MEXF 2090 cells for the indicated library plates under nutrient starvation and replication stress.
[D] Quantification of stress-specific fitness for the indicated KO populations of MEXF 2090 cells. Members of the same protein complex are color-coded.
Green and red lines indicate the z-score thresholds used to define populations with enhanced or decreased stress-specific fitness, respectively.
[E] Distribution of the expression levels of various epigenetic regulators assigned in the indicated sub-categories. Whiskers indicate the 10th and 90th percentiles, with outliers omitted for clarity. P-value from two-tailed Mann-Whitney test.

3.3.3 Epigenetic deregulation enhances cancer cell fitness under stress

3.3.3.1 Dichotomous effect of epigenetic deregulation on proliferation in unperturbed conditions

Initially, I explored how epigenetic deregulation affects the fitness of cancer cells growing in unperturbed conditions. 82% of the KO populations in melanoma cells displayed similar fitness to the negative controls, indicating that in most of the cases proliferation rates were unaffected [Figure 19]. Conversely, a shift in the distribution in proliferation rates was observed upon epigenetic deregulation in lung cancer cells, with more than 80 KO populations exhibiting growth retardation [Figure 19]. The above data reveal a dichotomous behavior where the effect of epigenetic deregulation on proliferation is dependent on the cancer type and probably on underlying global molecular characteristics of those cells (e.g. genomic instability). Similar experimental approaches need to be performed in additional cancer cell models that are characterized by diverse signalling, mutational and differentiation status to dissect the effect of epigenetic deregulation in unperturbed conditions. It is worth mentioning that during data processing, populations with severely decreased fitness in unperturbed conditions (like in the case of lung cancer cells) were discarded from the dataset to minimize potential inflation of the ratios during the calculation of the stress-specific phenotype.

Only a minimal fraction of either melanoma or lung cancer KO populations exhibited increased fitness suggesting that at least in these two cancer models interference with cell cycle checkpoints and enhanced proliferation are not broad consequences of loss-of-function mutations in core epigenetic regulators (Barnum & O'Connell, 2014).

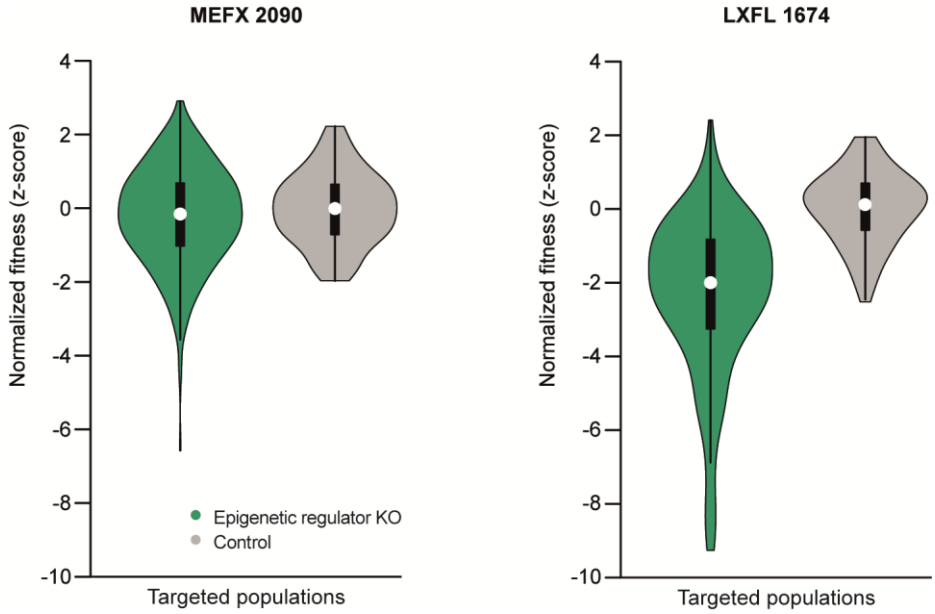


Figure 19: Fitness of epigenetically disrupted cancer cells in unperturbed conditions

Distributions of fitness from the large-scale assays in unperturbed conditions for the KO libraries of MEXF 2090 or LXFL 1674 cells. White circles show the medians; box limits indicate the 25th and 75th percentiles and whiskers extend 1.5 times the interquartile range from the 25th and 75th percentiles.

3.3.3.2 Numerous mutations in diverse epigenetic regulators promote cancer cell survival in nutrient starvation

Next, I went on to explore the stress-specific fitness. In melanoma cells, a broad increase in survival was observed upon epigenetic deregulation. Specifically, 91 KO populations exhibited increased fitness after three or seven days of growth under nutrient starvation [Figure 20A]. This behavior consisted of varying degrees of fitness advantage, ranging from top hits (e.g. KOs for either member of the PRC2 complex) to subtle responses that were consistently different from the negative controls (e.g. *HIST1H1B*-KO). Similar behaviors were also observed in the lung cancer cells, albeit to a lesser extent, where gene inactivation of 38 genes conferred an advantage to cells under stress [Figure 20B]. One possibility explaining the milder phenotype is the slow proliferation rate that characterizes the lung cancer cells along with the increased sensitivity of this model to mutations in epigenetic regulators. The detected broad increase in fitness is specific under stress, as disruption of the epigenetic machinery resulted in either neutral or decreased fitness in melanoma and lung cancer cells respectively [Figure 19]. As mentioned earlier both cancer models used in this study exhibit a maximal KO efficiency of approximately 70% [Figure 10]. The presence of a fraction of cells that remain unedited and the short time frame of the time points where the fitness was interrogated both suggest that the detected phenotypes are a potential underestimation of the actual effects imposed upon epigenetic deregulation. Follow-up experiments by Fabrizio Simeoni (Scaffidi

lab) in PDX models from colorectal, pancreatic and bladder carcinomas, demonstrated increased fitness of cancer cells upon inhibition of various epigenetic catalytic activities, further establishing confidence that the stress-resistance phenotype is not cancer type specific.

Notably, the favorable phenotypes were distributed across all examined layers of epigenetic regulation in both cancer models, indicating that the acquired resistance to stress is independent of specific catalytic activities and molecular functions [Figure 20C]. Comparison of the fitness behaviors of KO populations that were profiled in both cancer models revealed both shared and cancer type specific effects [Figure 20D-E]. Epigenetic regulators have variable expression patterns across different cancer models and possibly differential dependencies upon gene inactivation, thus differences are expected. An additional reason that could contribute to the observed differences, is the presence of some false negatives mainly in the slow-growing lung cancer cells.

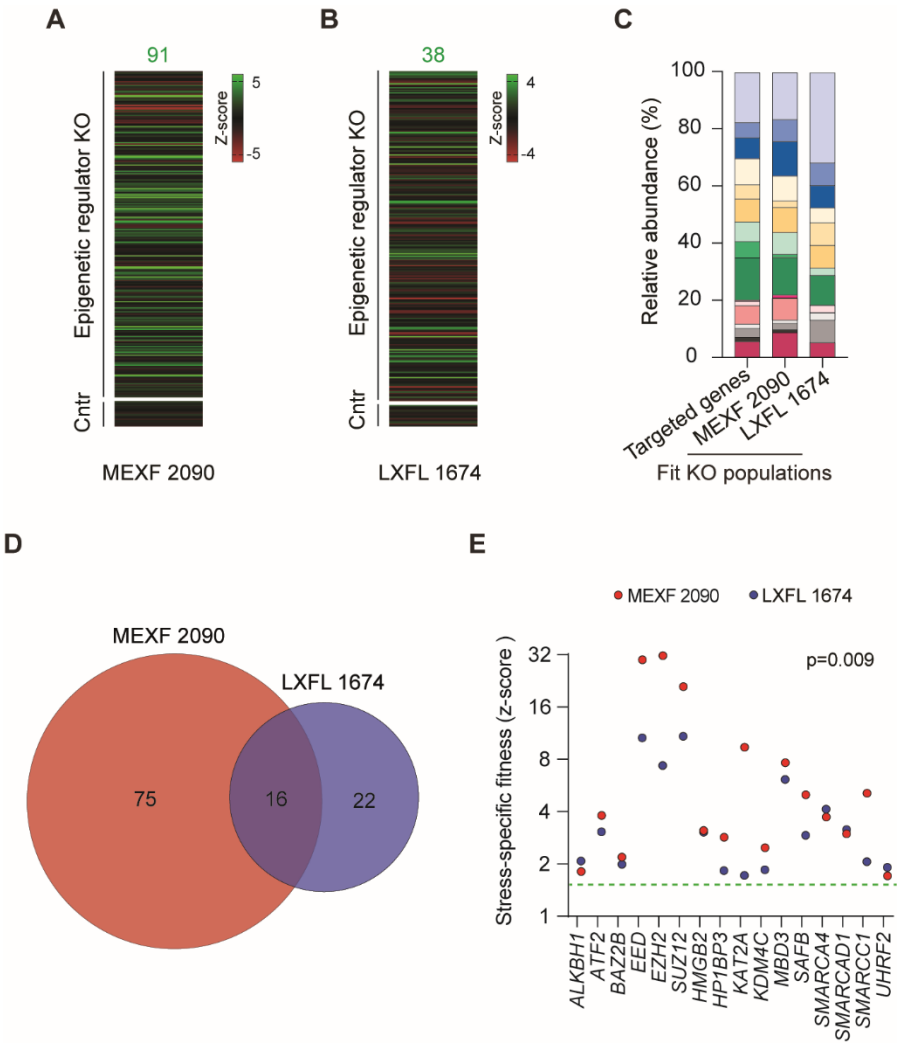


Figure 20: Widespread selective advantage of epigenetically disrupted cells under nutrient starvation

[A-B] Heatmap visualizing z-scores for the KO populations of MEXF 2090 (A) or LXFL 1674 (B) cells grown under nutrient starvation. KO populations in which non-expressed genes were targeted or that

exhibited reduced fitness under unperturbed conditions are not shown (see methods). The number of populations with enhanced fitness (green) is indicated.
[C] Representation of functional classes of epigenetic regulators among the genes exhibiting enhanced fitness in the indicated cell lines, compared to the representation among all targeted genes.
[D] Venn-diagram showing the number of KO populations that demonstrated fitness advantage in MEXF 2090 or LXFL 1674 cells.
[E] Quantification of stress-specific fitness for the indicated KO populations for both cancer models. The p-value of the Spearman correlation is shown

3.3.3.3 Robust validation of phenotypes from the large-scale fitness assay

Although the technical noise in our system is minimal and accounted for, it is still possible that some phenotypes were falsely classified as positive or negative in the large-scale fitness assays [Figure 18]. In line with that, it is worth reminding that the assessment of the fitness relationships was performed in fixed time points relatively early during the response to stress (3 and 7 days in starvation), where milder phenotypes can still fluctuate around the threshold of significance. To i) confirm the validity of the selected fitness threshold and the respective classification and ii) better understand the kinetics of response at longer time scales I selected 38 distinct populations (and 5 as controls) that exhibited survival advantage in the fitness assays. The selection of the different populations was done based on the following criteria: a) all KOs demonstrate increased survival under stress as assessed in the large-scale fitness assays, b) the extent of the phenotypic advantage varies and is inclusive of both top performers but also cases of subtle deviations from the controls. c) the selected panel of populations carries mutations in genes distributed across all the functional classes of the epigenetic network (Table 8). Along this thesis, except if stated otherwise, similar principles were followed for the selection of subsets of KO populations for specific experiments.

The validation fitness assay confirmed the robust detection of phenotypes, as at the later time points the majority of the KO populations demonstrated increased cell count compared to the negative controls [Figure 21A]. On top of that, monitoring those relationships over time revealed cases of shifting in the relative ranking among the examined populations, suggesting that early response is not an absolute predictor of the endpoint phenotype [Figure 21B]. For example, 2 days after exposure to stress *SIRT5*-KO cells exhibited a stronger resistance to stress compared to the ones where the chromatin remodeller SMARCD1 was inactivated. At the latest time point, this relationship was reversed with *SMARCD1*-KO cells being a top performer while the former population exhibited a much subtler phenotype. Such cases are suggestive of a complex response to stress where the final cell count of each population is the result of both the initial resistance to starvation but also the presence of potential compensatory mechanisms that emerge (and/ or are enriched for) during the later time points enabling the long-term survival of cells under chronic stress.

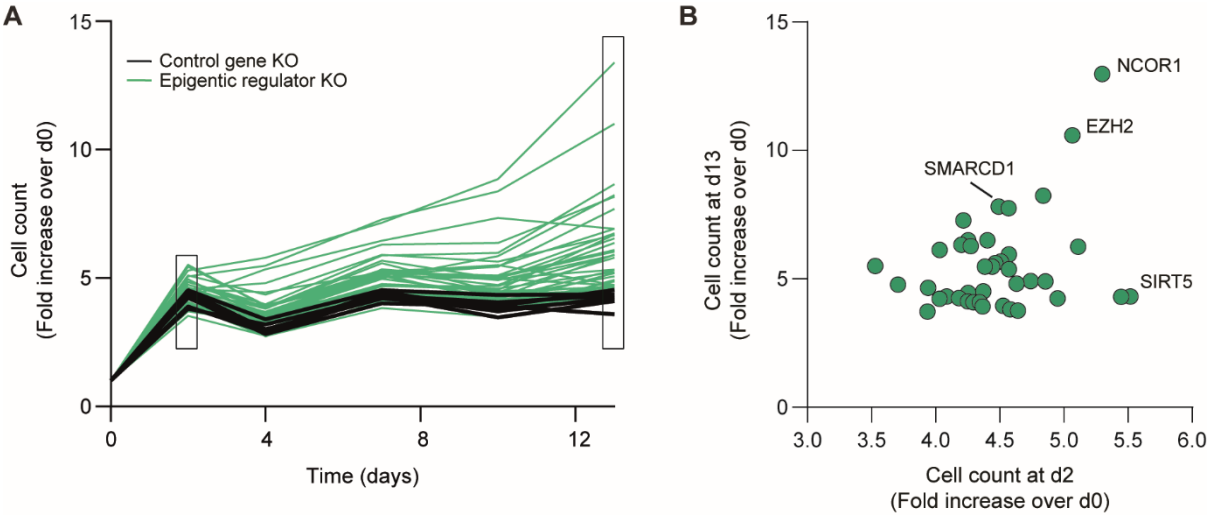


Figure 21:Validation of gene-stress relationships identified in the large-scale fitness assays

[A] Growth kinetics of 38 KO and 5 control populations of MEXF 2090 cells selected for validation of the results from the large-scale assay. Values represent mean from two biological replicates. Black boxes indicate the examined time points presented in B.
[B] Scatter plot of the relative cell count of epigenetically deregulated cells in the indicated time points. Each dot is a KO population.

3.3.3.4 Epigenetic deregulation increases the relative abundance of proliferating to apoptotic cells under nutrient starvation

To get a better understanding of how disruption of the epigenetic network benefits cancer cells under conditions of nutrient scarcity, I selected 17 KO populations and monitored their apoptosis and proliferation under glutamine starvation by probing Caspase-3 activity and EdU incorporation, respectively. In line with previous observations in the parental melanoma cells, three days after exposure to stress proliferation was severely halted in control populations accompanied by increased cell death. Notably, all epigenetically disrupted populations retained higher proliferation/apoptosis ratios, thus explaining their increased fitness in the early response to starvation [Figure 22B]. Some KO populations with similar cycling capacity under stress, demonstrated different levels of apoptosis (*MTA2*-KO and *MBD3*-KO cells). This suggests that the despite the overall antithetic trends in apoptosis and proliferation, these cellular behaviors are not necessarily entangled [Figure 22B]. Following those patterns over time, revealed a gradual increase in the detected differences, in line with the phenotypic deviation observed at later time points in the fitness assays [Figures 21A and 22A].

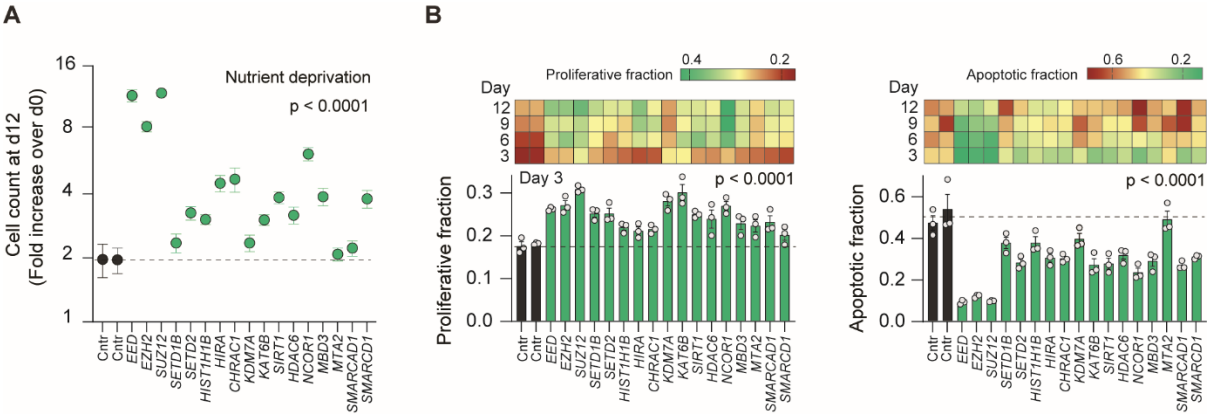


Figure 22: Epigenetically disrupted cells retain higher proliferation/apoptosis ratio during their immediate response to nutrient starvation

[A] Fitness assay comparing selected KO populations of MEXF 2090 cells after 12 days of growth under nutrient deprivation. The gene targeted in each population is indicated. Cntr: control population expressing sgRNAs targeting the non-expressed gene *TNP2*. Values represent mean ± SEM from three biological replicates. Two independent sets of control cells are shown, with the dashed line indicating their mean value. P-value from one-way ANOVA assessing overall differences among genotypes.
[B] Quantification of proliferative (left) and apoptotic (right) fraction in the indicated KO populations assessed in A, at the indicated time points. Values in the bar graphs represent mean ± SEM from three biological replicates. The dashed line indicates the average value of the two sets of control cells. P-value from one-way ANOVA. The mean value of the three replicates is visualized in the heatmaps. The heatmap colours are reversed for the two markers so that fit cells (high proliferative fraction and low apoptotic fraction) are green.

During the initial experiments exploring the parental response of melanoma and lung cancer cells under nutrient starvation, I observed the emergence of stress-resistant colonies after prolonged exposure to stress [Figure 11E]. To directly assess if epigenetic deregulation can alter the fraction of cells that can survive under stress, I selected a few representative KO populations and plated them sparsely followed by quantification of the surviving colonies after 12 days under starvation [Figure 23A]. All the examined KO populations showed a higher number of independent stress-resistant subpopulations compared to the control [Figure 23B]. Thus, the enhanced fitness detected at the population level reflects, to some extent, a higher fraction of cells that survive and continue to proliferate under stress. This correlation though is not absolute. In the large-scale assays and the validation experiments *HIRA*-KO cells exhibited a smaller survival advantage under starvation compared to *SMARCD1*-KO cells. This relationship was inverted in the clonogenic assays with more resistant colonies being present upon the inactivation of the chromatin remodeller SMARCD1. Although fitness advantage is overall correlated with the number of resistant colonies, the above observation suggests that additional factors contribute to the extent of the phenotypic advantage such as the emergence and selection of adaptive behaviors that can drive the proliferation of each colony in the long run (in line also with the observations in the validation fitness assays, where the ranking of fitness relationships changed over time, Figure 21).

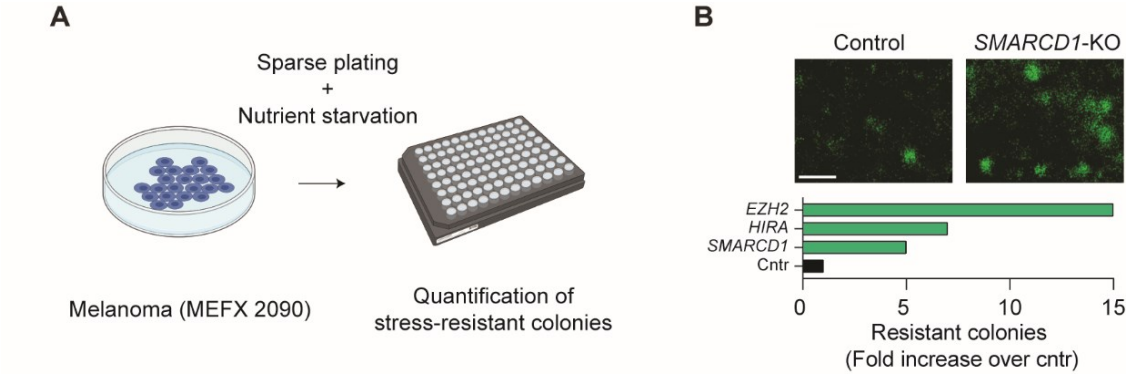


Figure 23: Increased stress-resistant colonies in epigenetically disrupted cells under nutrient starvation

[A] Schematic representation of the clonogenic assays.
[B] Representative images and quantification of stress-resistant subpopulations in the indicated KO populations after 12 days of growth under nutrient deprivation. Scale bar: 5 mm

Collectively, the data presented in the above subchapter (3.3.3.1-3.3.3.4) from the large-scale fitness assays and the subsequent experiments in selected panels of KO populations support the notion that the inactivation of numerous and diverse epigenetic regulators results in the acquisition of a stress-resistant phenotype that enables cancer cells to survive and proliferate under nutrient starvation.

3.3.3.5 Defining additional stress conditions for large-scale fitness assays in MEXF 2090 cells

One scenario that could explain the data presented so far, is that epigenetic deregulation leads to a specific metabolic rewiring that lowers the dependence of cancer cells on glutamine. A first indication against this scenario stems from the broad selective advantage observed in starvation, as it is unlikely that the disruption of highly diverse enzymatic activities would create the same dependency in a single amino acid [Figure 20A-C]. Nevertheless, if differential pre-existing sensitivity to glutamine starvation was underpinning the survival advantage, it is unlikely that the same cells would be able to survive a distinct type of stress condition that requires the activation of alternate molecular pathways during response and adaptation. Thus, I explored the response of cancer cells to additional stressors in order to identify specific conditions that could be used in fitness assays, similar to the ones presented so far. Having established that the phenotype is not dependent on the cancer type, I focused on MEXF 2090 cells for the subsequent analysis. In detail, varying densities of melanoma cells were cultured in diverse unfavorable contexts and growth kinetics were monitored over time. The examined

contexts mimic intracellular or environmental challenges that cancer cells commonly experience during tumor evolution.

Intracellular stress conditions

To induce either oxidative or replicative stress, I cultured melanoma cells under increasing concentrations of hydrogen peroxide (H_2O_2) and hydroxyurea (HU) respectively. The effect on cell fitness induced by H_2O_2 was significantly variable depending on the initial cell density indicating that the linearity of comparison is lost within a small window of initial technical variability. This can pose a significant challenge in large-scale experiments where such variability is an inherent aspect of the experimental setup. Additionally, exposure to H_2O_2 induced the formation of big cellular clumps that hindered the subsequent image acquisition and data analysis [Figure 24A]. Thus, oxidative stress was excluded from the subsequent assays. In contrast to H_2O_2, no variability was observed among different initial cell densities when cells were exposed to HU. Two concentrations of this stress agent were selected to induce replicative stress, resulting in 40% or 60% growth retardation respectively [Figure 24B].

Extracellular stress conditions

The tumor microenvironment is often characterized by a limited supply of nutrients accompanied by hypoxia and local acidification. To define acidic conditions suitable for downstream experiments, I exposed melanoma cells to increasing concentrations of hydrochloric acid (HCl). Similar to what was observed upon HU treatment, initial cell density did not affect the response of cells to acidosis. Two concentrations of HCl were selected, leading again to a 40% and 60% growth delay, respectively [Figure 24B]. Finally, I was unable to find a suitable O_2 concentration that affected the growth kinetics of melanoma cells. Cells cultured in decreasing concentrations of O_2, ranging from 1% to 0.1% combined with varying exposure times, propagated at identical rates compared to cells growing in unperturbed conditions. Hypoxia was thus not included among the stress conditions probed in subsequent experiments [Figure 24C].

Collectively, multiple different stress contexts were explored in MEXF 2090 cells, and two of them were selected for downstream large-scale fitness assays, namely environmental acidification and replicative stress.

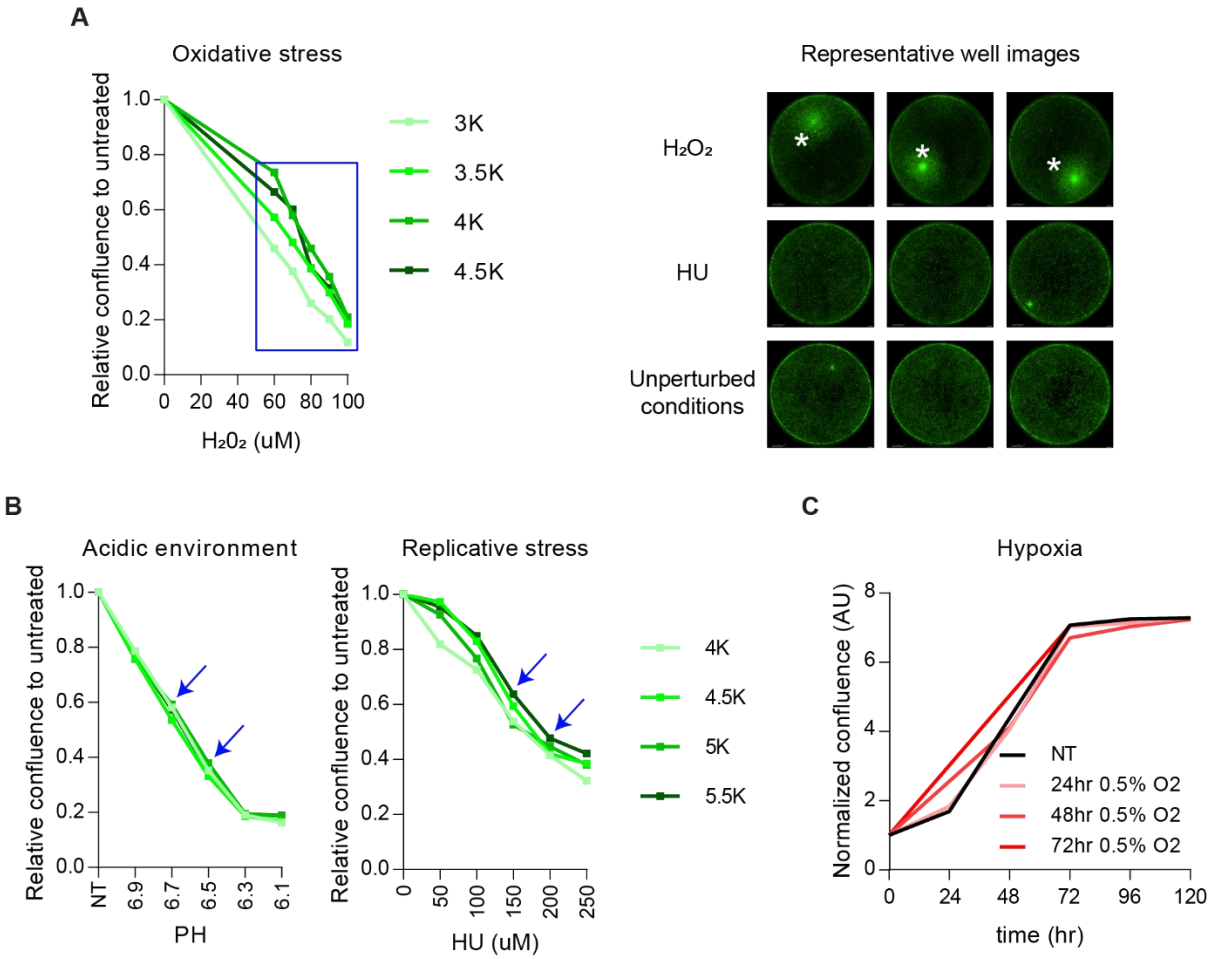


Figure 24: Titration of various stress conditions in MEXF 2090 cells

[A-B] Different densities of MEXF 2090 cells were seeded in 96-well plates. 24 hr later, increasing concentrations of stress agents were administered followed by time-lapse monitoring of their growth kinetics. The stress depth was estimated at the time point where the untreated cells reached plateau.
[C] Growth kinetics of cells growing in various hypoxic conditions, or in normoxia as a control.
Asterisk: big cellular clumps present when cells are exposed to H_2O_2.
Box: variable stress depth achieved in cells with different plating densities.
Arrow: the stress conditions that were selected for the subsequent large-scale assays.

3.3.3.6 Numerous mutations in diverse epigenetic regulators promote cancer cell survival in acidic conditions

Having explored the response of melanoma cells to additional stress conditions, I then asked if the broad stress-specific advantage upon epigenetic deregulation is restricted to nutrient starvation or shared across other conditions. To that end, I utilized the MEXF 2090 KO library and explored its response in acidic conditions, which have opposing effects on cellular metabolism compared to glutamine deprivation (Yoo et al., 2020). More specifically, environmental acidification rewires the metabolism of cancer cells from glycolysis to increased glutamine utilization, a switch that is dependent on HIF-2a activity (Corbet et al., 2014).

Comparable to nutrient starvation, 97 epigenetically disrupted cancer populations distributed across all functional classes demonstrated increased fitness [Figure 25A-B]. Validation experiments in a selected panel of KO populations confirmed the accurate phenotypic detection from the large-scale fitness assays [Figure 25C]. On top of that it confirmed the presence of KO populations with increased resistance to both nutrient deprivation and environmental acidification supporting the notion that epigenetic deregulation provides a general advantage to cancer cells to withstand stress in hostile settings [Figures 22A and 25C].

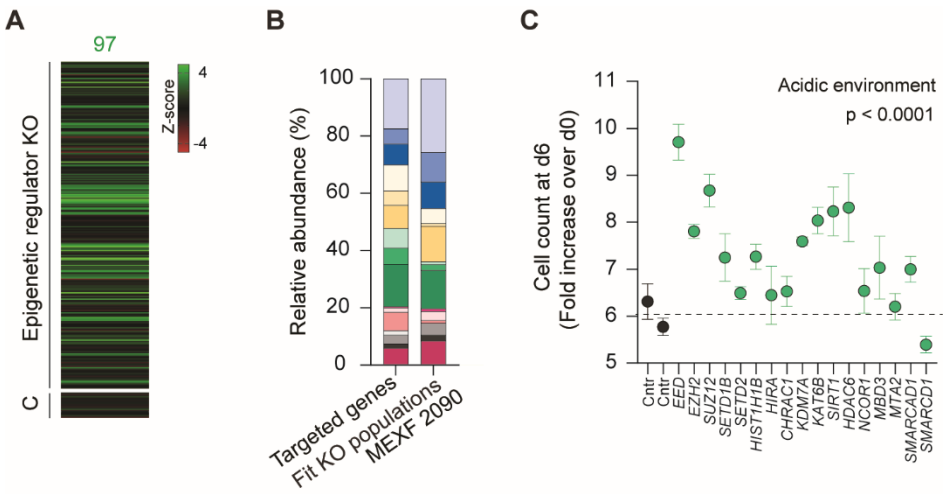


Figure 25: Widespread selective advantage of epigenetically disrupted cells in acidic conditions

[A] Heatmap visualizing z-scores for 262 KO populations of MEXF 2090 cells grown in acidic environment. KO populations in which non-expressed genes were targeted or that exhibited reduced fitness under unperturbed conditions are not shown (see methods). The number of populations with enhanced fitness (green) is indicated.
[B] Representation of functional classes of epigenetic regulators among the genes exhibiting enhanced fitness in the indicated cell line, compared to the representation among all targeted genes.
[C] Fitness assay comparing the indicated KO populations of MEXF 2090 cells after 6 days of growth in acidic conditions. Cells were split at d3 after reaching confluence. Values represent mean ± SEM from three biological replicates. Two independent sets of control cells are shown, with the dashed line

indicating their average value. P-value from one-way ANOVA assessing overall differences among genotypes.

The phenotype of the MEXF 2090 KO library was also assessed in a third unfavorable condition, where cells were challenged by hydroxyurea treatment leading to replicative stress. Inactivation of epigenetic regulators only had minor effects on the response to this type of internal stress (median z-score: 0.24 vs 0.84 for nutrient deprivation and 0.86 for acidity), suggesting that disruption of epigenetic control mainly increases the ability of cells to cope with cell-extrinsic challenges (Table 8).

The data obtained so far demonstrate that damaging mutations in a wide range of chromatin and DNA regulators enhance the survival of cancer cells in distinct hostile conditions, independently of the tissue of origin. Nutrient deprivation and environmental acidification exert antithetic effects on cellular metabolism; thus, it is not expected that survival in these conditions will be conferred by the same pathways. The broad survival advantage detected in both conditions is indicative of a more universal underlying mechanism. This is corroborated also by the significant numbers of fit KO populations, as it is highly unlikely that more than 90 distinct KO populations would acquire the same specific metabolic or transcriptomic rewiring towards survival. On top of that, the observation that the genes conferring survival advantage, upon their inactivation, belong to all examined epigenetic layers, confirms that the broad survival advantage is not dependent on specific epigenetic activities (e.g. methylation or acetylation of histones, etc). Altogether the above observations are indicative of a network-level effect, where multiple and distinct perturbations within the epigenetic regulatory layer, converge towards a common underlying mechanism of resistance to stress, that can be advantageous in diverse hostile settings.

3.3.3.7 Competitive advantage of epigenetically disrupted cells in stressful TMEs

In vitro competition assay

The established fitness relationships presented so far were identified through *in vitro* experiments where each population grows independently. Thus, I asked if similar phenotypic relationships can be detected in epigenetically disrupted populations in the presence of competing control cells. To that end, I generated two MEXF 2090 cell lines, stably expressing either GFP or mCherry (see also Methods section 2.10.1). Then I generated a panel of 18 KO populations in diverse epigenetic proteins via lentiviral transduction of the respective targeting sgRNAs (similar pools to the ones used for the large-scale fitness assays). After generating the KO lines, I co-cultured the epigenetically disrupted mCherry populations with control GFP

cells under glutamine starvation. To account for potential differences in the growth of the parental fluorescent cell lines, I also included as a control a co-culture of wild-type cells labelled with GFP and mCherry.

After a 12-day growth period under nutrient deprivation, the majority of the epigenetically deregulated cells exhibited positive selection, at varying degrees in line with the observations from the large-scale fitness assays [Figure 26A-B]. The simultaneous monitoring of the fluorophores allows for the estimation of the normalized growth of each population in the competing conditions. In the case of competing control-GFP and control-mCherry cells, it is expected that the cells will get equally stressed in response to nutrient starvation. Thus, the normalized growth for either fluorophore should be similar. If an epigenetically disrupted population exhibits fitness advantage through a cell-intrinsic mechanism (i.e. that does not affect the conditions for the surrounding cells) this should result in an increased mCherry normalized growth while the GFP growth remains unaltered (or slightly decreased if there is indirect competition for limiting resources). On the other hand, if the fitness advantage is conferred through a non-cell-autonomous mechanism (e.g. secretion of pro-survival molecules) then the expectation is that the normalized growth of the control-GFP cells will also be indirectly enhanced [Figure 26C]. Inspection of the relative growth of either fluorophore revealed a variable response, with the majority of the KO populations exhibiting a fitness advantage that is in line with cell-intrinsic underlying properties [Figure 26D]. In contrast to the tight clustering of the control replicates (black), most of the replicates within the epigenetically disrupted cancer cells exhibited significant variability. A notable example is *EED*-KO cells, where the replicates demonstrated drastically different behaviors in line with either non-cell autonomous or cell-intrinsic modes of survival under nutrient starvation. Similar behaviors were also detected in other epigenetically deregulated populations, like KO cells for *CHRAC1, HDAC6, MTA2 or SETD2* [Figure 26D]. This observation explains the significant deviation observed in the fraction of mCherry+ cells in Figure 26B.

It is intriguing to speculate that disruption of the epigenetic function enhances fitness via a universal underlying mechanism that promotes resistance to stress in the short term thus providing cells with critical time to acquire distinct secondary traits that increase the potential for long-term survival in stress. This scenario predicts the stochastic emergence of different cellular behaviors and is in line with the variable phenotypes in replicate populations observed in the competition experiment.

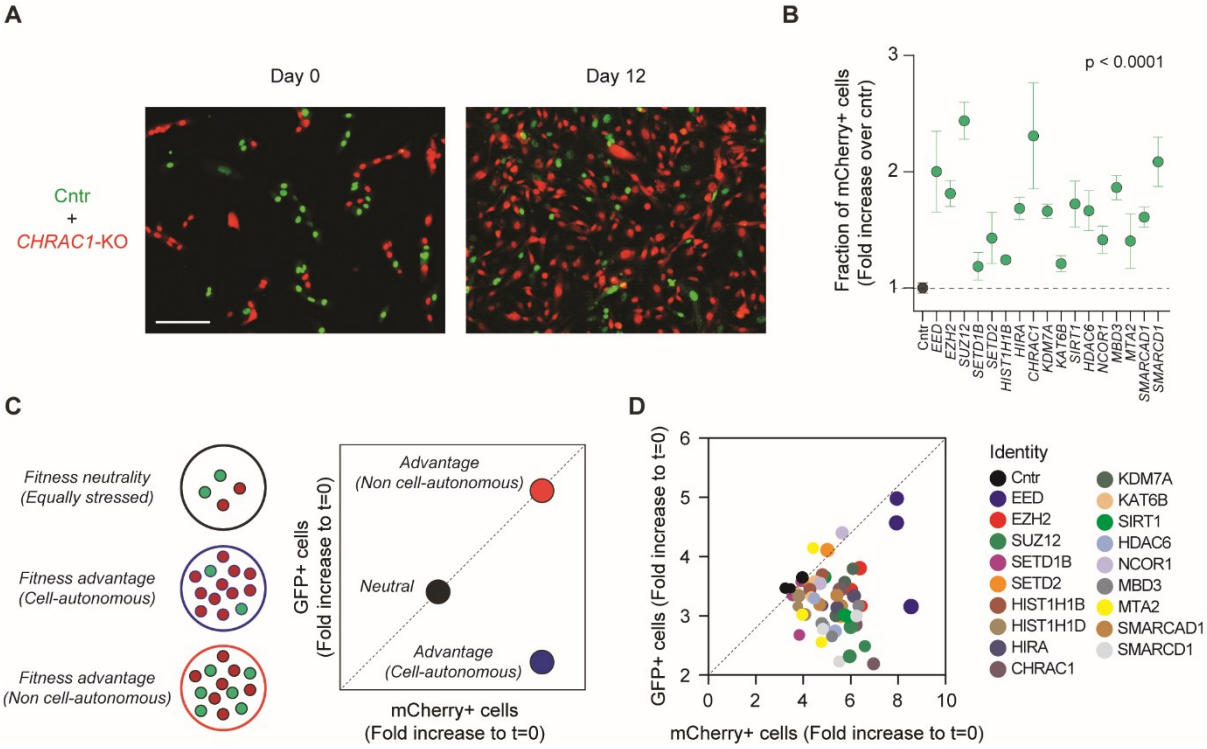


Figure 26: Competitive advantage of epigenetically disrupted cells under nutrient starvation

[A-B] Representative images (A) and quantification (B) of *in vitro* competition assays in which equal numbers of mCherry-labelled KO and GFP-labelled control populations of MEXF 2090 cells were seeded in the same wells and grown under nutrient deprivation for 12 days. Values represent mean ± SEM from three biological replicates. P-value from one-way ANOVA assessing overall differences among genotypes. Cntr indicates co-cultures in which mCherry- and GFP-labelled control cells were mixed. Scale bar: 200 µm.
[C] Schematic illustrating the potential models underlying the fitness relationships between competing populations.
[D] Scatter plot of the normalized growth of control GFP+ or mCherry+ KO populations. Each dot represents a different replicate within the examined populations. Three different replicates are included in each comparison. The different colours represent the pair-wise competitions.

In vivo competition assay

Note: Intradermal injections of cancer cells were performed with the help of Cristina Morales Torres (Scaffidi lab). Subsequent injections and handling of the mice were carried out by the Biological Research Facility at the Francis Crick Institute. Next-generation sequencing was performed by the Advanced Sequencing Facility and processing of the generated raw data by Harshil Patel. Subsequent analysis was performed by Ioannis Loukas. Embedding of tumors in paraffin, subsequent sectioning and H&E staining was performed by the Experimental Histopathology Facility at the Francis Crick Institute. Imaging was performed by Ioannis Loukas.

The stress-dependent selective advantage of epigenetically deregulated cells was detected in *in vitro* experimental settings. However, within a growing tumor many additional factors can shape the fitness landscape, thus necessitating further exploration of the identified relationships in the in vivo setting. Aiming to assess if similar phenotypic behaviors can be observed in evolving tumors, I selected *EZH2*-KO cells and performed *in vivo* competition assays. To do so, I injected equal fractions of control and *EZH2*-KO cells into NOD.Cg-Prkdc^{scid}Il2rg^{tm1Wjl}/SzJ (NSG) mice and allowed induced tumors to grow for approximately five weeks. Naturally occurring tumors are often characterized by hostile environmental properties, like limited supply of nutrients and oxygen, due to the impaired architecture and functionality of their vasculature. To further deteriorate the properties within the tumor microenvironment, I treated established cancers (approximately three weeks after the injection) with Bevacizumab, an antibody against VEGF-A that inhibits angiogenesis [Figure 27A] (Ferrara, 2005). At endpoint I extracted the tumor specimens from the mice and subsequently quantified the relative fraction of control or *EZH2*-deficient cells via next-generation sequencing of the retrieved sgRNAs (*TNP2*-targeting or *EZH2*-targeting guides integrated into controls cells and KO cells) [Figure 27A]. To confirm the success of the anti0angiogenic treatment, I stained tumor sections with an antibody against phosphorylated S6 ribosomal protein (pS6), an indicator of mTOR activity that is dependent on the levels of available nutrients. Indeed, treated tumors were characterized by low levels of this marker, confirming the emergence of tumors with a hostile setting, that is suitable for exploring the potential phenotypic differences between control and EZH2-KO cells [Figure 27B].

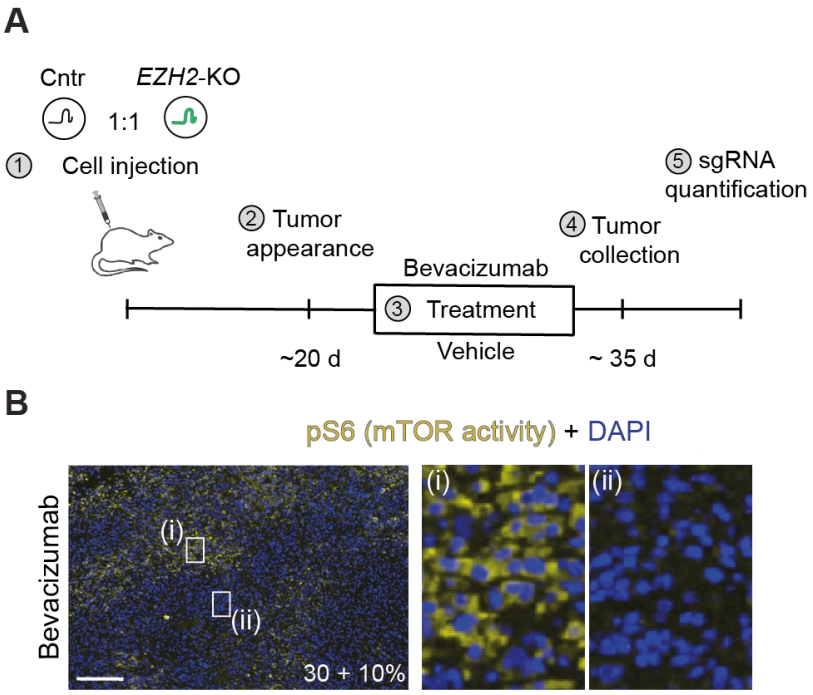


Figure 27: *In vivo* competition assay with manipulation of the tumor microenvironment

[A] Schematic illustrating the experimental pipeline followed for the *in vivo* competition assay. Circles represent MEXF 2090 cells with integrated sgRNAs, which were used to quantify the relative abundance of the two subsets of cells in tumors.
[B] Immunofluorescence microscopy section of a tumor treated with Bevacizumab. Magnified versions of the boxed regions are shown on the right. The percentage of $pS6^{high}$ tumor area is indicated (mean ± SEM from four regions).

In agreement with the results from the *in vitro* assays, *EZH2*-KO cells consistently outcompeted control populations in all Bevacizumab-treated tumors, with some samples consisting almost exclusively of *EZH2*-KO cells [Figure 28A]. Immunostaining of tumor sections for H3K27me3, a histone modification that is mediated by the PRC2 complex, revealed complete loss of the mark in bevacizumab-treated specimens confirming the selection of EZH2-deficient cells in evolving tumors [Figure 28B].

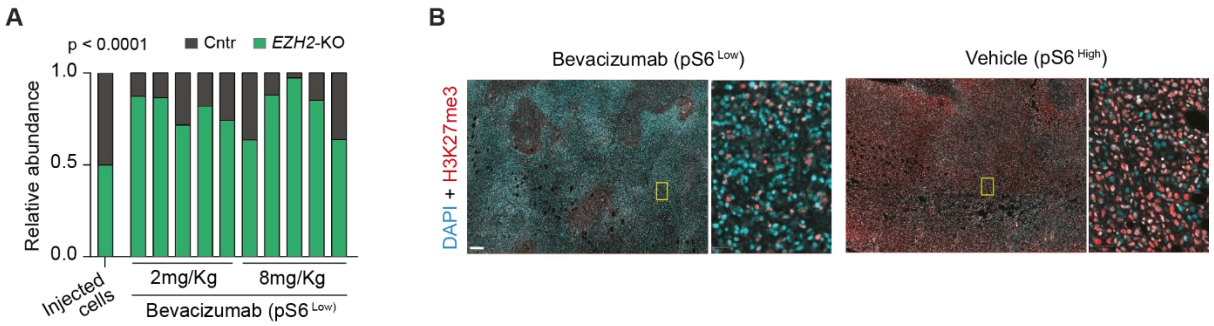


Figure 28: Stress-dependent selection of *EZH2*-KO cells in evolving tumors

[A] Quantification of the changes in the fraction of *EZH2*-KO cells in the Bevacizumab treated tumors relative to the injected cells, as assessed by next-generation sequencing of the sgRNAs. P-value comparing the number of *EZH2*-KO enriched tumors in Bevacizumab- and vehicle-treated tumors from two-tailed χ2 test.
[B] Immunofluorescence microscopy of tumors treated with the indicated substances. Depletion of H3K27me3 in nuclei identifies *EZH2*-KO cells. Boxed regions are shown as magnified. Scale bars, 200 mm.

Contrary to the above homogeneous pattern, unperturbed tumors demonstrated a dichotomous behavior. In half of the cases *EZH2*-KO cells were enriched at the expense of the control population, while in the rest of the cases the detected fractions were comparable [Figure 29A]. One possible explanation for this heterogeneity is that the examined tumors are characterized by varying basal levels of nutrient availability and environmental acidification, thus creating different selective pressures on the competing populations. Indeed immunostaining for pS6, revealed that the vehicle-treated tumors that consist of equal fractions of both control and KO populations are characterized by high levels of pS6, indicative of a nutrient-rich tumor mass, whereas the ones occupied predominantly by epigenetically disrupted cells showed reduced levels of this marker [Figure 29B].

This observation from the *in vivo* setting corroborates the findings from the large-scale fitness assays where EZH2 inhibition resulted in increased fitness only in the presence of stress.

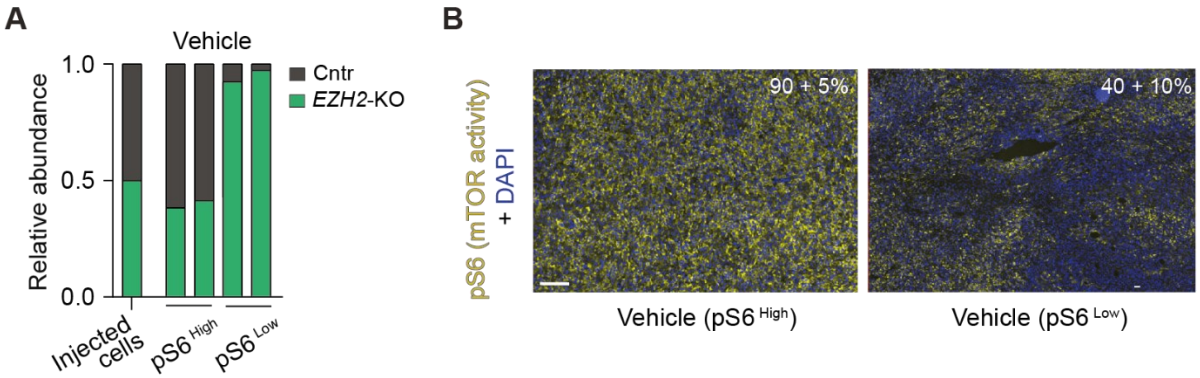


Figure 29: Basal levels of nutrient availability dictate the selective advantage of *EZH2*-KO cells in unperturbed tumors

[A] Quantification of the changes in the fraction of *EZH2*-KO cells in the vehicle treated tumors relative to the injected cells, as assessed by next-generation sequencing of the sgRNAs.
[B] Immunofluorescence microscopy of vehicle treated tumors. The percentage of $pS6^{high}$ tumor area is indicated (mean ± SEM from four regions).

Overall, the above data demonstrate that disruption of epigenetic control provides a competitive advantage under nutrient starvation. Initial evidence from *in vivo* models suggests that this notion can hold true in the more complex setting of evolving tumors. Further experimental investigation is required to explore the extent of the functional link between epigenetic deregulation and increased survival of cancer cells *in vivo* (See section 5.2.3 for further discussion on the matter).

3.4 Conclusion

The scope of this chapter was to assess if disruption of the epigenetic network can alter the fitness of cancer cells under unfavorable conditions. To that end, I selected two distinct cancer models, melanoma and NSCLC cells, that are devoid of nonsense mutations in epigenetic regulators and explored their response to various stress conditions that are relevant for cancer cells. I then optimized the steps towards the successful completion of large-scale fitness assays where I systematically inactivated hundreds of epigenetic-related genes in both cancer cell models, mimicking the frequent loss-of-function mutations observed in cancer patients, and assessed cell fitness under diverse environmental conditions. Finally, selected *in vitro* and *in vivo* experiments confirmed the validity of the detected fitness relationships and elucidated how epigenetic deregulation affects specific aspects of the cellular response to stress [Figure 30].

The key data presented in this chapter are summarized below [Figure 30]:

- ✓ Selection of two distinct cancer models with a functional epigenetic regulatory network, and characterization of their response to nutrient starvation, a hallmark of the tumor microenvironment [Figures 9 and 11].
- ✓ Cancer cells exhibit functional heterogeneity in response to limiting nutrients, a key feature of tumor microenvironments. In this setting, favorable phenotypes can spontaneously emerge in the absence of underlying genetic events [Figures 12 and 13].
- ✓ Systematic disruption of more than 300 core epigenetic regulators revealed widespread survival advantage of KO populations under nutrient starvation or environmental acidification [Figures 20 and 25].
- ✓ The stress-resistant phenotype is cancer-type independent and is conferred by inactivating mutations in multiple genes, distributed across all layers of epigenetic regulation [Figures 20 and 25].
- ✓ In depth analysis of cellular behaviors under nutrient starvation revealed that disruption of diverse epigenetic proteins converges in a cellular state that enables cells to resist stress-induced cell death and retain higher proliferation thus increasing the probability of long-term survival, through the selection of pre-existing or *de novo* acquisition of favorable traits [Figures 21, 22, 23 and 26].
- ✓ Manipulation of the tumor microenvironment *in vivo,* provided initial evidence for the competitive stress-dependent selection of epigenetically disrupted cancer cells in evolving tumors [Figures 27, 28 and 29].

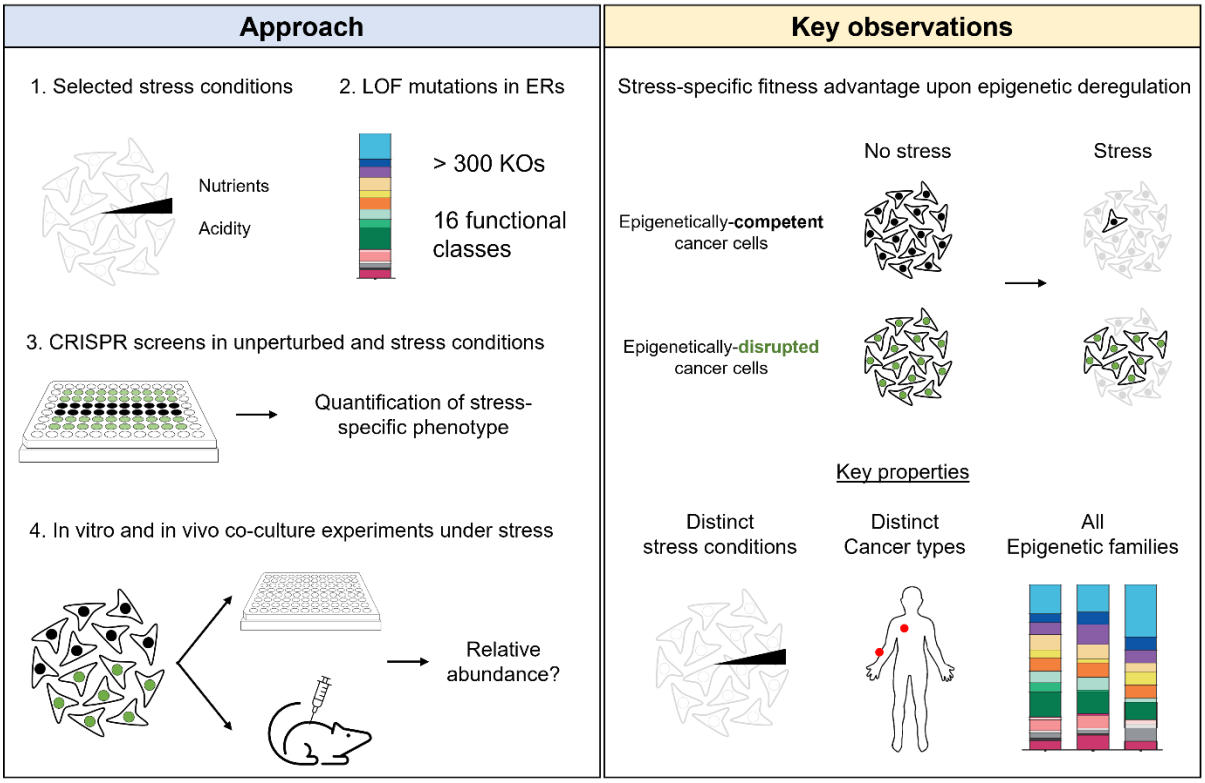


Figure 30: Summary of the experimental methodologies and key observations reported in Chapter 3

Collectively, the above data establish a link between epigenetic deregulation and increased fitness of cancer cells under stress. A plethora of questions emerges from the above observations. To name a few:

- How are the qualitative (functional classification of fit KOs) and quantitative (number of fit KOs) aspects of the broad resistance affected by distinct driver mutations, signalling pathways and other systematic factors?
- Is the stress-resistant phenotype shared in other extracellular conditions that are not directly related to cellular metabolism? Could it be relevant for resistance to therapy?
- Are there contexts where disruption of the epigenetic regulatory network is detrimental?

Exploration of such questions along with future experimental directions to tackle them are presented in detail in Discussion section 5.2. The most interesting question though is the identification of the underlying mechanism that mediates the resistant phenotype of KO cells. Is there a shared molecular outcome of epigenetic deregulation that enables cells to survive in diverse stress conditions? In the next chapter of my thesis, I am extensively exploring various candidates trying to answer this exact question.

Chapter 4.
Phenotypic inertia underpins the stress resistance of epigenetically disrupted cells

4.1 Introduction

The data presented in Chapter 3 of my thesis revealed a link between disruption of the epigenetic network and enhanced fitness of cancer cells under unfavorable conditions. This advantageous phenotype could be the result of various underlying cellular mechanisms [Figure 31].

a) Genetic diversity

Epigenetic mechanisms can affect events within the genome in various and diverse ways (Shen & Laird, 2013). For example, the faithful repair of genetic lesions, mediated by the DNA Damage Response (DDR) pathway, is dependent on tightly regulated alterations in the chromatin state (Sulli, Di Micco & d'Adda di Fagagna, 2012). Considering this functional link between genome and epigenome, one possibility worth exploring is that epigenetic deregulation increases the generation of genetic abnormalities and/or hinders the fidelity of their resolution, thus ultimately leading to increased mutational burden and subsequently genetic and subclonal diversity. As a result, cancer cells have an increased probability to acquire an advantageous mutation that confers resistance to stress.

b) Transcriptional diversity (Bet-hedging)

Epigenetic mechanisms act as a regulatory layer promoting selective and faithful transcription in various contexts. Disruption of this regulation may lead to promiscuous transcription that increases the chance of expressing gene sets that could provide a survival advantage under specific stress conditions. Indeed, the emergence of rare subpopulations with distinctive transcriptional output has been reported to be important in the adaptation of cancer cells to targeted therapy (Shaffer et al., 2017). Conceptually, this mechanism is the equivalent at the transcriptional level of the genetic diversity model discussed above, with the major difference residing in the reversibility of such alterations (genetic vs non-genetic events).

c) Phenotypic plasticity

From a developmental perspective, epigenetic regulators have been implicated in defining cellular identity by establishing the faithful expression of specific transcriptional programs.

Disruption of the epigenetic network may lead to a more permissive chromatin state that enhances the capacity of cells to transit between distinct states in response to their environmental stimuli (Flavahan, Gaskell & Bernstein, 2017). In the context of stress resistance, which is the focus of this thesis, such fluidity in the cellular states could manifest as cells being able to more robustly (i) acquire an alternative cellular state that promotes survival in the respective stressor or (ii) fluctuate back to the initial state after the exposure to stress, thus lowering the probability of cell death [Figure 31].

It is worth mentioning that there is significant overlap between the transcriptional diversity and the plasticity models. This primarily stems from the fact that a more "plastic" cellular population can, but not necessarily will, provide a more heterogenous gene expression output. Along this thesis, I employ diverse experimental setups to dissect among these two scenarios.

d) Phenotypic inertia

Epigenetically disrupted cells could be unable to sense or respond to alterations in their stressful environment and subsequently promote an apoptotic program leading to cell death. Thus, their ability to tolerate stress in the short term may allow for increased survival in case of transient challenges or provide a critical time window for the acquisition of secondary adaptive behaviors under chronic stress.

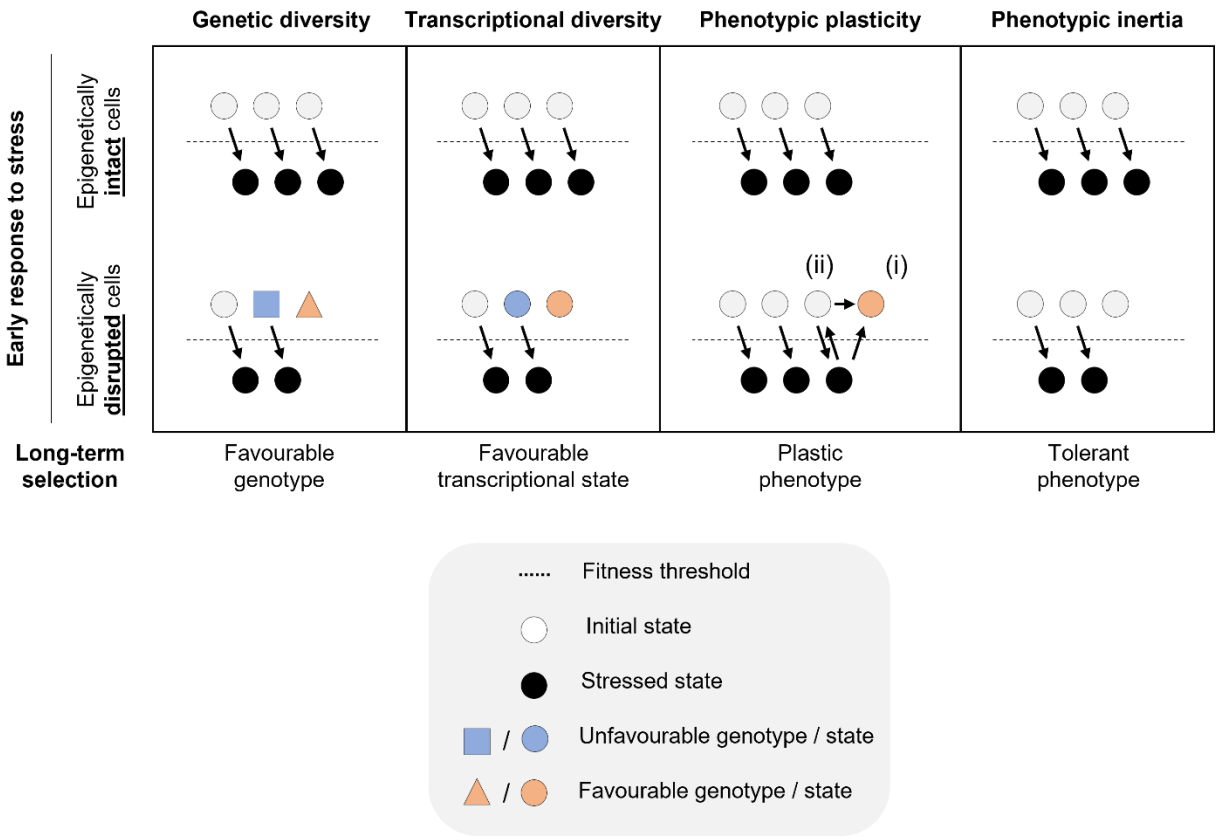


Figure 31: Distinct models of resistance to stress in epigenetically disrupted cancer cells

Alternative models of stress-resistance in epigenetically disrupted cells. For phenotypic plasticity, (i) indicates the transition to a transcriptionally rewired cellular state before or after reaching a stressed state and (ii) indicates restoration of the initial state.

4.2 Aim

In Chapter 4, I describe the steps taken to investigate the aforementioned models mediating the increased resistance of epigenetically disrupted cancer cells under stress. Based on the predictions created by each mechanism, I employ a variety of methodologies ranging from reversible drug administration to live-cell imaging and single-cell transcriptomics to dissect among them. Initially, I present data that interrogate the reversibility of the stress-resistant phenotype. Then, I directly assess cell state transitions in response to stress by employing a FRET based biosensor. Next, I describe the response of cancer cells to nutrient starvation and assess how epigenetically deregulated cells act differently at the transcriptional level. I go on to interrogate multiple aspects of the system like sub-population heterogeneity, gene expression variance and global alterations in transcriptional burst kinetics and assess how these parameters are affected by stress and mutations in epigenetic regulators. Finally, I propose that phenotypic inertia underlies the stress-dependent fitness advantage of epigenetically disrupted cells.

Note: Several data presented in this chapter are already published (Loukas et al., 2023). *Selective advantage of epigenetically disrupted cancer cells via phenotypic inertia* © 2023 by Ioannis Loukas et al is licensed under CC BY 4.0. The researchers that contributed to this chapter are acknowledged accordingly at the beginning of the respective subchapters.

4.3 Results

4.3.1 Interrogating the reversibility of the stress-resistant phenotype

4.3.1.1 The stress-specific selective advantage of epigenetically disrupted cells is not genetically encoded

To begin to dissect among the possible mechanisms underlying the fitness advantage of epigenetically disrupted cells, I first examined whether their behavior has a genetic basis [Figure 31]. Several lines of evidence argue against the contribution of genetic abnormalities to the stress-resistant phenotype. As stated in Chapter 3, the cellular models used in this study are clonal lines derived from PDXs, thus the pre-existing diversity at the genomic level is minimal. Additionally, after generating the KO populations, I avoided prolonged culture of the cells to minimize potential divergence (genetic and non-genetic). Finally, in the large-scale fitness assays the phenotype of various KO populations was assessed after short exposure to environmental challenges (3 days in acidic conditions, 3 and 7 days in nutrient starvation). This narrow timeframe is in discordance with a model where *de novo* acquisition and selection of rare favorable genetic events underlies the observed stress resistance.

Aside from these initial indications I went on to directly assess the genetic diversity model. Let's assume that disruption of the epigenetic landscape promotes genomic instability and subsequently the generation of subclonal populations with favorable mutations. In that case, reverting to the original epigenetic state after the initial perturbation should not affect the presence of the genetically distinct subpopulations and thus the detected survival advantage. To explore this dependency, I utilized the chemical compounds that were previously shown to promote cancer cell survival to nutrient starvation [Figure 14] and I performed a wash out experiment. More specifically, pre-treated melanoma cells were cultured in nutrient starvation and monitored for 9 days, confirming the advantageous behavior upon inhibition of diverse epigenetic functions [Figure 32]. Subsequently, the stress-resistant populations were recovered in media containing L-glutamine either in the presence or absence of the drugs followed by a second round of growth in starvation. Notably, in all cases the emerged resistance to stress was completely lost upon drug withdrawal, suggesting that the detected phenotype does not have a genetic basis [Figure 32].

Altogether, the above data directly demonstrate that the stress-specific survival advantage of epigenetically disrupted cancer cells is not genetically encoded. It is worth noting that this does not mean that epigenetic deregulation cannot lead to genomic instability, but rather implies that there are contexts, like the one examined here, and time scales during cancer evolution where alternative mechanisms can mediate the advantageous behaviors emerging upon epigenetic deregulation.

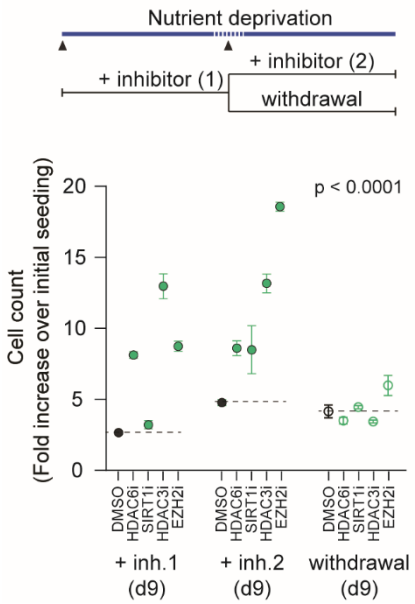


Figure 32: The stress resistant phenotype of epigenetically disrupted cells is not genetically encoded

Fitness assay using MEXF 2090 cells treated with the indicated epigenetic inhibitors and grown under nutrient starvation following the protocol indicated above. Cells were grown for 9 days in the presence of inhibitors, split and grown for an additional 9 days in the presence or absence of the inhibitors. At each seeding (arrowhead), cells were pleated at identical density. Values represent mean ± SEM from three biological replicates measured at the time point indicated in brackets. P-value from two-way ANOVA assessing the effect of genotype in the comparison + inhibitor vs withdrawal at d18. The p-value of the differences between matching samples in the two conditions is $p < 0.01$ for all KO populations and non-significant for DMSO.

4.3.1.2 Increased survival of epigenetically disrupted cells in fluctuating environmental conditions

Having provided evidence against the genetic diversity model, I then examined if enhanced cellular plasticity could explain the stress-resistant phenotype. A first prediction of the plasticity model is that cells should be able to reversibly transit between distinct cellular states in response to their fluctuating environmental conditions. To test this, I exposed a panel of diverse KO populations to multiple rounds of nutrient starvation alternated with growth in unperturbed conditions. Monitoring the cellular behavior over time revealed that epigenetically disrupted cells are not locked in the favorable state dictated by starvation but can rather

reversibly transit between states in response to their environmental stimuli [Figure 33A]. Interestingly, the fitness of all examined populations in the second exposure to starvation was increased, suggesting that cells may retain a form of molecular memory that entails a more robust response during the second exposure to the stressor.

A second prediction of the plasticity model is that cells should have the capacity to alter their cellular behavior continuously thus being able to adapt to distinct sequential stress conditions. After challenging the same panel of KO populations in nutrient starvation, the cells were recovered and cultured in acidic conditions, another relevant stress for cancer cells (Wei et al., 2020). At the second challenge, all populations were seeded at identical numbers to allow for a robust comparison of the fitness in acidic conditions, irrespective of any fitness differences in nutrient starvation. All the examined epigenetically disrupted populations demonstrated increased capacity to tune their response based on the stress they experience and subsequently achieved enhanced fitness compared to control[5] cells in both stress conditions [Figure 33B].

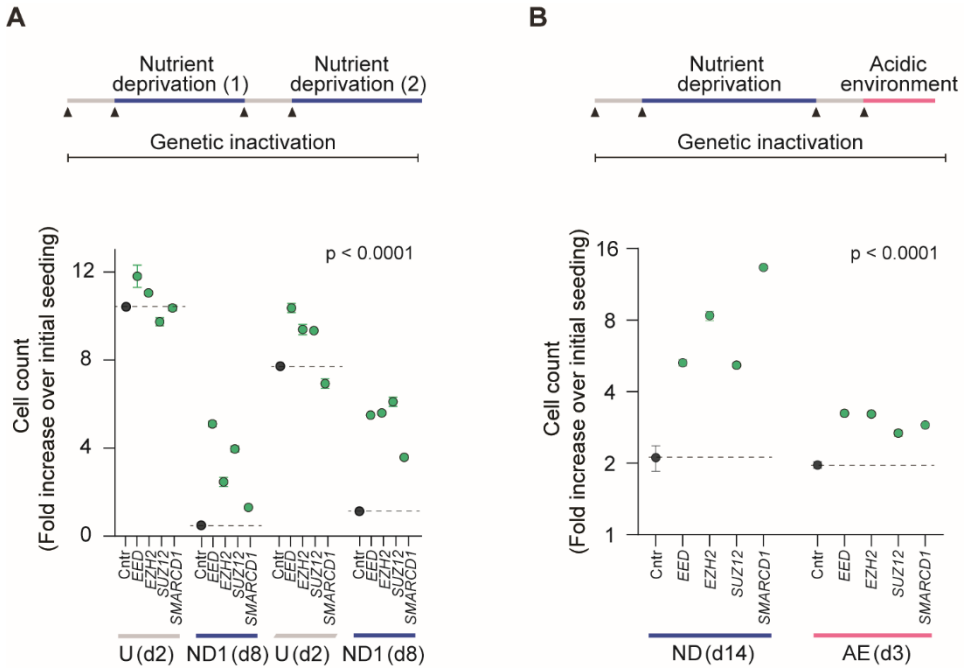


Figure 33: Epigenetically disrutped cells exhibit increased survival to fluctuating environmental conditions

Fitness assay using the indicated KO populations of MEXF 2090 cells grown following the protocols indicated above. Cells were grown for 2 days in unperturbed conditions (U), for 8 (A) or 14 (B) days under nutrient starvation (ND) and for 3 days in acidic environment (AE). The endpoint of each treatment was determined by when the fittest population approached confluence and depended on how severely cell growth was affected in each condition. At each environmental change, cells were seeded at identical density as the initial seeding (arrowheads). Values represent mean ± SEM from three biological replicates measured at the time point indicated in brackets. P-value from two-way ANOVA assessing

[5] Throughout this Chapter "control" cells describe MEXF 2090 populations that are either KO for *TNP2* (non-expressed gene) or are treated with DMSO, depending if the experimental approach involves genetic ablation or pharmacological inhibition of epigenetic regulators respectively.

the effect of genotype across conditions. The p-value of the differences between matching samples in two consecutive conditions is $p < 0.01$ for all populations, indicating that cells adjust their behavior to fluctuating environments.

The ability of epigenetically disrupted cancer populations to survive within fluctuating environmental conditions (even when normalizing the effect of the first stressor) is a strong indicator of phenotypic switching and is in line with the predictions made by the plasticity model. In an alternative scenario, if epigenetically disrupted cells are characterized by an increased tolerance to stress, they would still be able to more robustly survive after both the first and the second exposure to hostile conditions. However, the advantage in this case is not provided by the ability of cells to switch their phenotypic state and adapt to the stress (plasticity), but rather through the selection of a stress-tolerant state (inertia) [Figure 31]. This observation provides an interesting example where conceptually distinct mechanisms that rely on antithetic cellular behaviors can result in the same phenotypic outcome (increased survival) and cannot be distinguished by approaches that solely rely on fitness quantification. Thus, additional experimental methodologies are required to discriminate between these two alternative mechanisms that may underpin the fitness advantage of epigenetically disrupted cells.

4.3.2 Disruption of epigenetic control does not promote cell state transitions

4.3.2.1 Characterizing the transcriptional response to nutrient starvation

Note: Bulk RNA sequencing was performed by the Advanced Sequencing Facility at the Francis Crick Institute. Processing of the generated raw data was performed by Harshil Patel. Subsequent analysis was performed by Ioannis Loukas.

Considering that phenotypically inert and plastic cells could not be distinguished by their behavior in the fitness assays, I explored other aspects of the system that could discriminate between the two models. If epigenetically deregulated cells are characterized by increased plasticity (or transcriptional heterogeneity), it is expected that they will adopt a favorable expression state after exposure to stress, which will gradually be enriched by positive selection. By comparing the transcriptomic profiles of control and KO cells after prolonged culture under nutrient starvation, I speculated that such “adaptive” signatures would be detectable. On top of that, comparing the transcriptomes of cells in unperturbed or nutrient-deprived conditions would also reveal the overall gene signatures that are responsive to stress. Such stress-responsive genes could be then used to directly assess state transitions

relevant to my system and further ask how epigenetic deregulation can affect this phenomenon. Finally, by examining the transcriptional alterations induced by inactivating diverse epigenetic regulators (in the absence of stress), I would explore the presence of pre-existing differences in gene expression that could prime these populations for a better response to stress.

I began by characterizing the transcriptional alterations, using 3'mRNAseq, induced by nutrient starvation in both control MEXF 2090 melanoma cells and in a selected panel of KO populations. Similar to previous experiments presented in this thesis, the selection of KO populations was based on the following criteria: i) selection of KO populations for genes belonging to distinct epigenetic functional classes, ii) selection of KO populations with varying extent of survival advantage under stress, as determined in the respective fitness assays and iii) the targeted genes are frequently inactivated in cancer.

More specifically, I genetically inactivated the chromatin remodeller SMARCD1 (Carlson & Laurent, 1994), the linker histone HIST1H1B – an integral component of chromatin (Scaffidi, 2016) and the three core subunits of the Polycomb repressive complex 2 (PRC2), namely EZH2, EED and SUZ12 (Piunti & Shilatifard, 2021). The PRC2 members and SMARCD1 have antithetic effects on gene expression leading to transcriptional attenuation or activation respectively. In line with this notion, inactivation of any PRC2 subunit resulted predominantly in gene upregulation in unperturbed conditions whereas the opposite effect was observed in the case of *SMARCD1*-KO cells [Figure 34A]. The inclusion of all three PRC2 members served as a means to assess the biological significance and validity of the observed patterns since its two structural components (EED and SUZ12) were expected to have nearly identical phenotypic effects. Indeed, a striking similarity was observed between *EED*- and *SUZ12*-KO populations, with only a trivial number of transcriptional changes detected between them [Figure 34B]. The phenotype observed in *EZH2*-KO cells was highly similar but diverged from the rest of the complex components, which can potentially be attributed to the reported non-canonical functions of EZH2 in cancer cells (catalytic activity independent of PRC2 complex) (Huang et al., 2021).

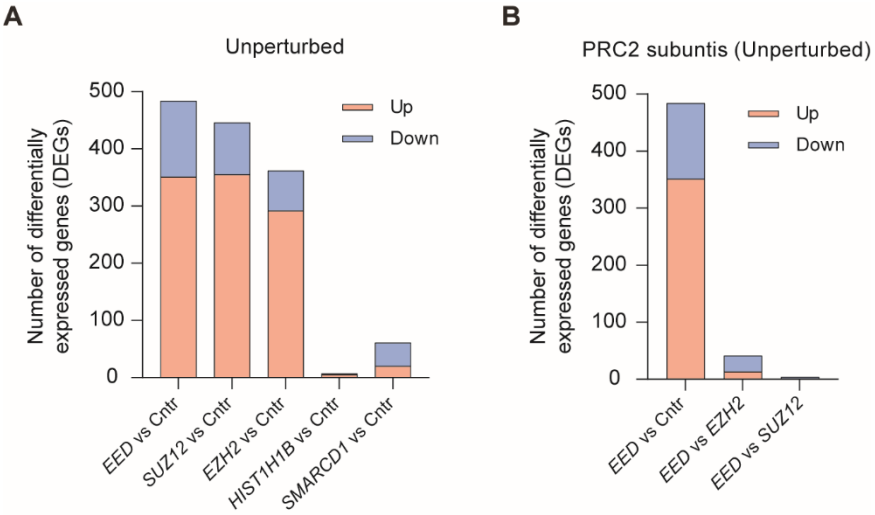


Figure 34: Differentially expressed genes among KO populations in unperturbed conditions

[A] Number of upregulated (orange) or downregulated (blue) genes in the indicated pair-wise comparisons among the KO populations.
[B] Number of upregulated (orange) or downregulated (blue) genes in the indicated pair-wise comparisons among the PRC2 subunits.
A threshold of FDR<0.01 was used to define differentially expressed genes.

For each population I compared the global transcriptomic profile of cells before the application of stress (unperturbed conditions, d0) and after 12 days of growth in nutrient starvation (d12), where all the KO populations demonstrate varying degrees of survival advantage [Figure 35].

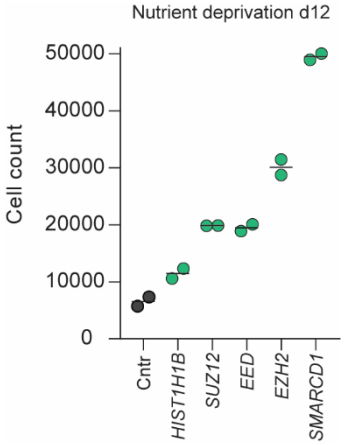


Figure 35: Fitness under nutrient starvation of the indicated KO populations of MEXF 2090 cells selected for bulk RNA-seq analysis

Hierarchical clustering of the analyzed populations revealed robust grouping of the biological replicates confirming the good technical quality of the generated dataset [Figure 36]. A dichotomous segregation of the KO populations was observed between samples cultured in

unperturbed conditions and under nutrient deprivation, suggesting that the challenge imposed by the stress and the subsequent transcriptomic alterations are dominant over KO-specific effects. Finally, the transcriptomic profiles from the PRC2 members diverged from the rest of the populations in the unperturbed conditions in line with the significant gene expression alterations that were detected in the pair-wise comparisons [Figure 34A]. Those alterations occurred predominantly in developmentally relevant genes, as expected by earlier studies of these proteins, and are irrelevant to the gene sets that are responsive to stress in the system under examination (data not shown).

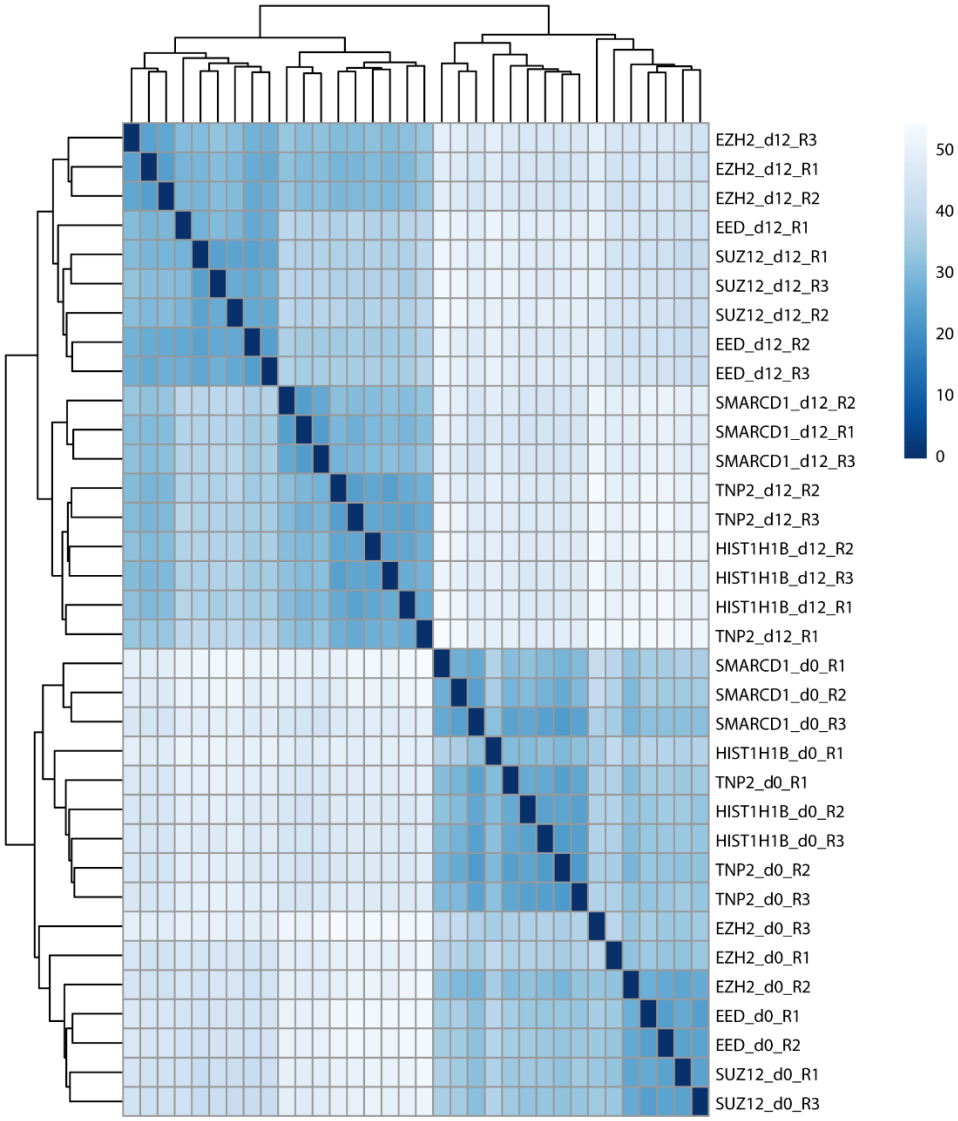


Figure 36: Hierarchical clustering of the assessed populations in unperturbed conditions or under nutrient starvation

The colour scale reveals the similarity among the different samples. For each KO population 3 distinct replicates (R1, R2, R3) are presented in two different conditions, unperturbed (d0) or under nutrient starvation (d12). Note that the stress treatment dictates the segregation of the samples.

4.3.2.2 Identification of stress responsive pathways

The transcriptomic response to nutrient starvation involved the differential regulation of thousands of genes [Figure 37A]. The downregulated genes are highly expressed and are significantly enriched in gene signatures related to the fitness of cancer cells, including cMYC target genes, cell-cycle regulators, genes involved in mTORC1 signalling and components of the oxidative phosphorylation pathway that acts as the predominant source of energy within the cells [Figure 37B-C]. At the same time, pathways associated with apoptosis and signatures associated with stress such as the NFkb target genes were significantly upregulated [Figure 37C]. On top of that, nutrient deprivation resulted in a strong inflammatory response, in line with previous reports on other cancer models, strengthening the confidence in the generated dataset (Gameiro & Struhl, 2018). From here on, I refer to the above identified signatures collectively as *fitness* or *stress* genes[6]. Wherever subdivisions of those gene sets are used this is indicated.

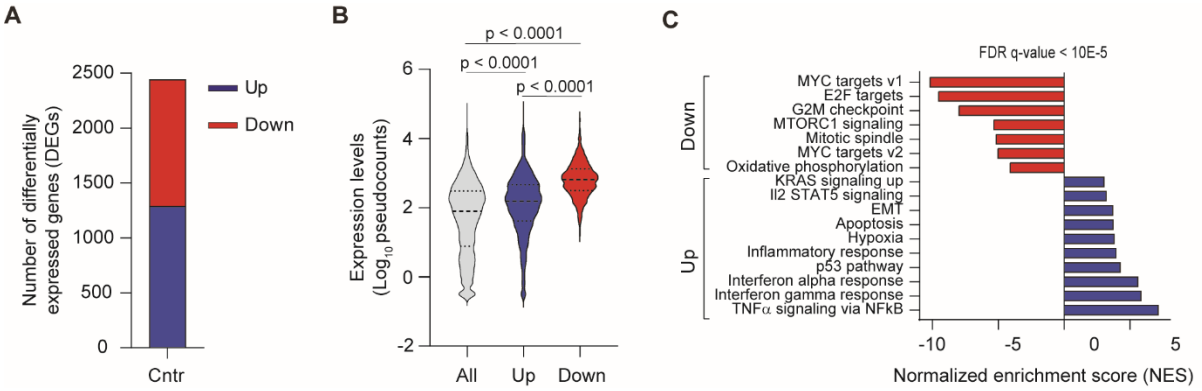


Figure 37: Identification of stress responsive pathways

[A] Number of upregulated (red) or downregulated (green) genes in control MEXF 2090 cells after 12 days of growth under nutrient starvation. A threshold of FDR<0.01 was used to define differentially expressed genes.
[B] Distribution of expression levels of the indicated groups of genes in unperturbed conditions. P-value from two-tailed Mann–Whitney U-test.
[C] GSEA analysis of stress-responsive genes (d12 vs d0 DEGs) identified in control cells.

The aforementioned transcriptomic alterations were identified in the context of nutrient starvation. To assess if these changes could be relevant for other unfavorable conditions, I cultured control MEXF 2090 cells in an acidic environment and assessed after 2 days the

[6] The classification of a gene as *fitness* or *stress* is specific to this study and was defined by the alteration the expression levels after exposure of melanoma cells to nutrient starvation. Generally, *fitness* genes are highly expressed in unperturbed cells and are involved in various aspects of proliferation, metabolism, biomass and energy production. On the other hand, *stress* genes are upregulated in unfavourable conditions and are associated with apoptosis, inflammation, p53 signalling, etc.

expression of key fitness and stress genes compared to their levels before the application of stress. Similar to nutrient starvation, representative stress genes underwent significant upregulation while the expression of fitness genes was attenuated, suggesting that diverse triggers can lead to a common cellular phenotype that characterizes cancer cells experiencing stress [Figure 38]. It is worth mentioning that similar alterations in grow-related and stress genes have been observed in cancer cells exposed to various extrinsic and intrinsic stressors (Emran et al., 2018; Gameiro & Struhl, 2018)

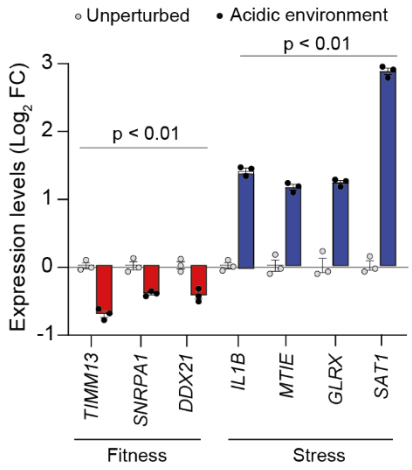


Figure 38: Differential regulation of fitness and stress related genes in response to environment acidification

qRT-PCR quantifying the changes in expression levels for the indicated stress and fitness genes in control MEXF 2090 cells after 2 days of growth in acidic environment. Values are shown as relative to the levels detected in unperturbed conditions. P-values from two-way ANOVA comparing the group of stress or fitness genes in unperturbed or treated cells.

I then explored the behavior of the epigenetically deregulated cells. Analysis of their transcriptomic response did not detect any adaptive signatures (at the bulk level) that could mediate their stress-dependent fitness advantage. On the contrary, similar patterns of transcriptomic changes in fitness and stress signatures were observed in the epigenetically disrupted cells after 12 days under nutrient starvation [Figure 39A-B]. However, the extent of these alterations was milder in the KO populations and was inversely correlated with the degree of survival advantage observed in the fitness assays [Figure 35]. Examination of the basal levels (day 0) of stress-responsive genes failed to detect any differences prior to stress application among the distinct populations, ruling out the possibility that the observed milder response could be attributed to pre-existing differences in those pathways [Figure 39C].

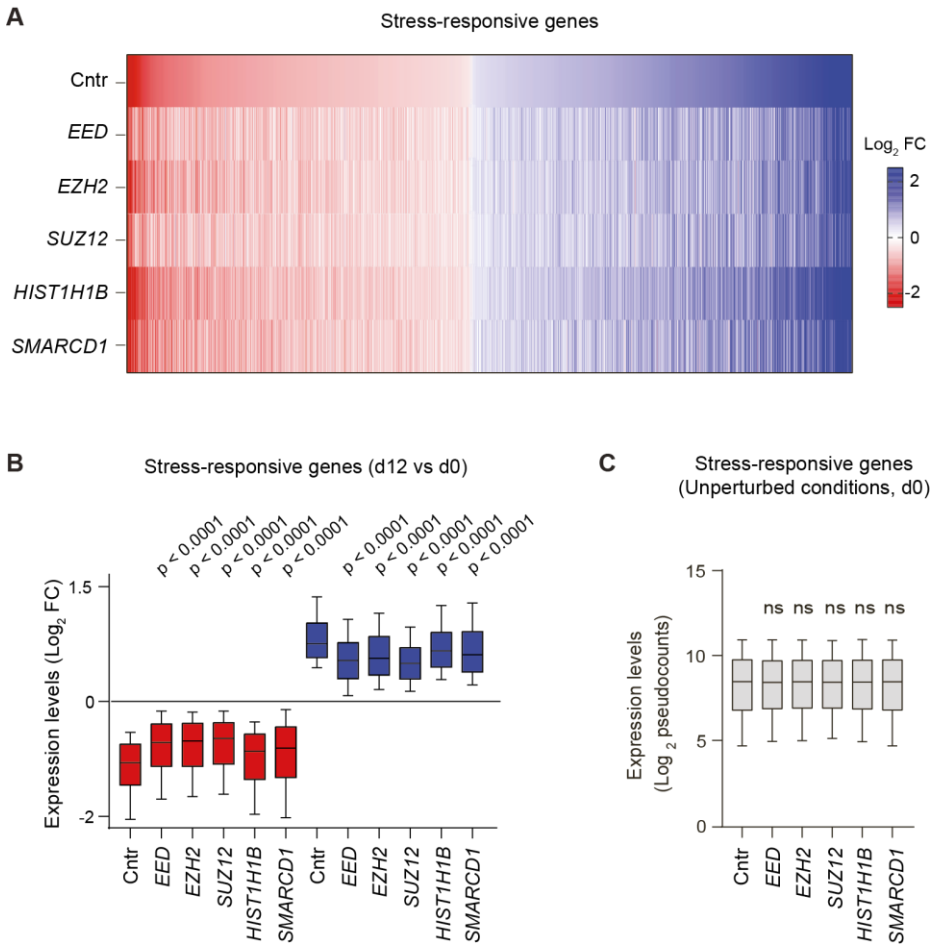


Figure 39: Epigenetically disrupted cells exhibit milder alterations to the stress responsive pathways

[A] Heatmap visualizing the changes in expression levels for stress-responsive genes defined in control MEXF 2090 cells, in the indicated cell populations.
[B-C] Distribution of the fold-changes (d12 vs d0) (B) or expression levels in unperturbed conditions (d0) (C) for stress-responsive genes defined in control cells. Whiskers of boxplots indicate the 10th and 90th percentiles, with outliers omitted for clarity. P-value relative to control cells from one-way ANOVA followed by Dunnett's test.

In summary, the above data indicate that epigenetically disrupted cancer cells acquire a less stressed phenotype in response to nutrient starvation, in line with the phenotypic deviation observed in the fitness assays. This initial interrogation of the transcriptomic response at the population level (bulk analysis) did not detect the activation of favorable transcriptional programs indicative of a rewired state that would be able to promote survival in nutrient shortage. Such alterations would include activation of alternate compensatory metabolic pathways, utilization of micropinocytosis or other forms of reported adaptation of cancer cells to L-glutamine starvation (Zhang, Pavlova & Thompson, 2017).

4.3.2.3 Assessing cell state transitions via a FRET biosensor

Note: Live-cell imaging was performed in collaboration with Colin D.H. Ratcliffe (Sahai Lab, The Francis Crick Institute). Subsequent image analysis and quantification were performed by Ioannis Loukas.

The absence of adaptive signatures from the bulk transcriptomics experiment provided a first line of evidence against the plasticity and the bet-hedging models. As stated earlier one possibility that could characterize the behavior of plastic cells, is their enhanced capacity to fluctuate back to the initial state after exposure to stress which in turn lowers the probability of cell death [Figure 31]. The identification of gene sets that are responsive to stress provides a framework for studying cell state transitions that are relevant for survival in our system. In particular, the transcriptomic analysis revealed a significant downregulation of genes involved in mitochondrial respiration (Oxidative phosphorylation; OXPHOS) in response to nutrient deprivation [Figure 37C]. To examine such transitions within the OXPHOS state I utilized a FRET-based biosensor that can monitor the metabolic state of living cells (Kondo et al., 2021). Monitoring of the OXPHOS state in response to stress in control and KO populations can uncover if the survival advantage of the latter is associated with increased fluctuations of stressed cells back to a non-stressed state (plasticity) or increased resistance to stress-induced alterations (inertia).

How does the FRET biosensor work? The sensor carries two distinct fluorophores with spectral overlap, namely eCFP (donor) and Citrine (acceptor). In the absence of glucose, excitation of the system results in emission of signal from eCFP. However, when intracellular glucose increases it induces a conformational change to the sensor that brings the two fluorophores in proximity, leading to energy transfer between them and the subsequent emission of fluorescence from Citrine at a different wavelength (FRET signal) [Figure 40A]. By quantifying the relative ratio of the FRET to eCFP signal, indirect but robust estimation of the oxidative phosphorylation activity (OXPHOS) of living cells can be achieved (Kondo et al., 2021). To generate a cell line stably expressing the FRET sensor, I transfected MEXF 2090 cells with plasmids encoding a) the glucose FRET biosensor and b) the PiggyBack transposase that mediates the stable integration of the sensor into the genomic DNA. 7 days post transfection, the melanoma cells were FACS sorted based on Citrine expression and a viable stable line was obtained [Figure 40B-C].

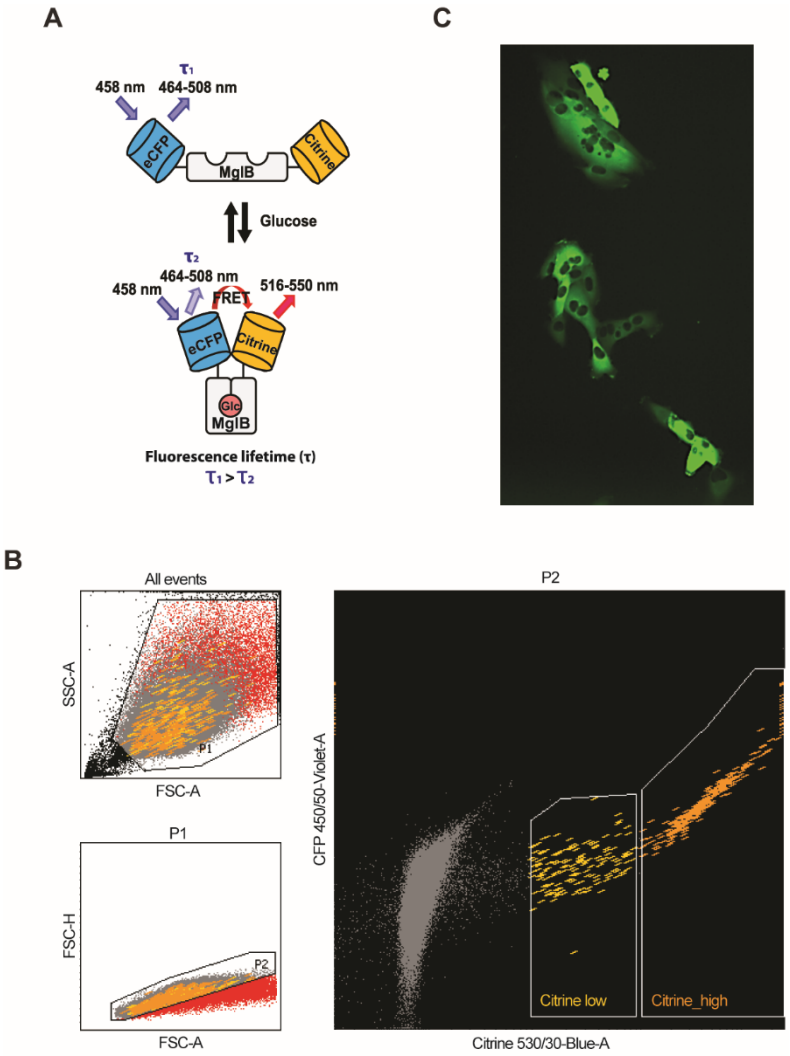


Figure 40: A FRET-based biosensor to monitor metabolic state transitions

[A] Schematic representation of the FRET biosensor used to monitor the metabolic state of cells. Adapted from (Kondo et al., 2021)
[B] FACS plot of the gates used to isolate a MEXF 2090 cell line stably expressing the FRET biosensor.
[C] Live-cell imaging of the MEXF 2090 cell line stably expressing the FRET biosensor.

The quantification of the metabolic state is done by measuring the fluorescence from perinuclear foci in individual cells [Figure 41A]. The FRET signals obtained by multiple foci in the same cell were highly correlated, suggesting that monitoring a single area per cell is sufficient to robustly infer the metabolic state of a cell [Figure 41A]. Additionally, no correlation was observed between the total expression levels of the biosensor and the detected FRET signal, indicating that there is no bias in quantifying the phenotype among cells that are characterized by heterogeneous expression of the sensor [Figure 41B]. To confirm that this FRET biosensor can be an indicator of the OXPHOS state within the cells, I cultured pre-treated DMSO and EZH2-inhibited cells under nutrient starvation and examined after 24 hours

both the FRET signal and the level of oxidative phosphorylation, as inferred by TMRE[7] staining. In line with what Kondo and colleagues previously reported in other cancer models, high FRET signal anticorrelated with the levels of oxidative phosphorylation, confirming the usability of this biosensor in our system [Figure 41C-D] (Kondo et al., 2021). Overall, I successfully generated a MEXF 2090 cell line stably expressing a FRET biosensor that can be utilized to monitor the metabolic state of living cancer cells.

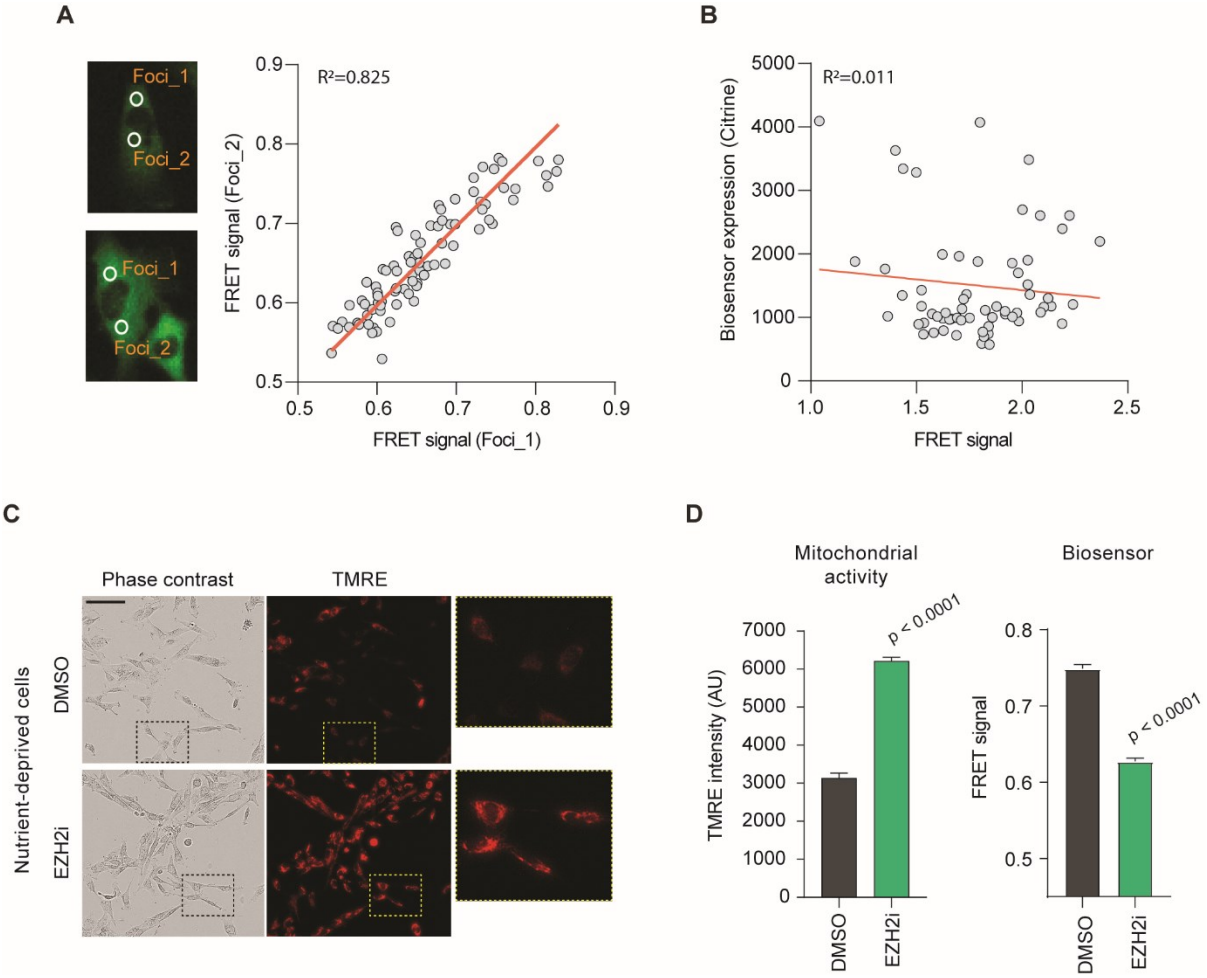


Figure 41: The FRET-based biosensor can robustly infer the metabolic state within living cells

[A] Scatter plot of the FRET signal obtained from 2 perinuclear foci in each cell. R-squared correlation from linear regression is shown.
[B] Scatter plot of the FRET signal and the expression level of the biosensor. Each dot reflects the quantifications obtained from a single cell. R-squared correlation from linear regression is shown.
[C] Imaging of MEXF 2090 cells treated with the indicated substances and grown under nutrient starvation for 24 hours. Mitochondrial activity is visualized by TMRE staining. A higher magnification of the regions boxed in yellow is shown on the right. Scale bar: 50 µm
[D] Quantification of TMRE staining and FRET signal in MEXF 2090 cells treated with the indicated substances. TMRE integrated intensity values represent mean ± SEM from 9 fields. FRET signal values represent mean ± SEM from 137 and 330 cells. P-value from two-tailed Student's t-test.

[7] Tetramethylrhodamine, ethyl ester (TMRE) is a cell-permeant, positively-charged dye that accumulates in active mitochondria due to their negative charge

Having optimized the technical aspects of the generated stable cell line, I went on to compare the relative OXPHOS levels in epigenetically deregulated and control cells. As described earlier the level of CRISPR-Cas9 mediated KO efficiency achieved in MEXF 2090 cells is ~70% [Figure 10]. To avoid the presence of unedited cells that could hinder the quantification at the single cell level, a chemical compound against EZH2 was used that leads to homogeneous inhibition of its function (EZH2i). Thus, I pre-treated cells with EZH2i or DMSO as a control and monitored their metabolic state during their initial response to starvation. Depriving cells of glutamine resulted in a significant decrease in OXPHOS levels that characterizes the stressed phenotype. On the other hand, EZH2i cells displayed milder alterations in response to starvation in line with the data obtained from the bulk transcriptomic analysis [Figure 42, also refer to Figure 37 and 39]. It is worth noting that in the absence of stress there was a subtle difference detected in the FRET ratio of the examined populations. However, follow-up experiments by Marta Milan (Scaffidi Lab) that directly examined OXPHOS levels in these populations in unperturbed conditions did not reveal any pre-existing differences, suggesting that the detected difference in the FRET ratios could be an artifact emerging from differences in the initial cell density or other technical factors (Loukas et al., 2023).

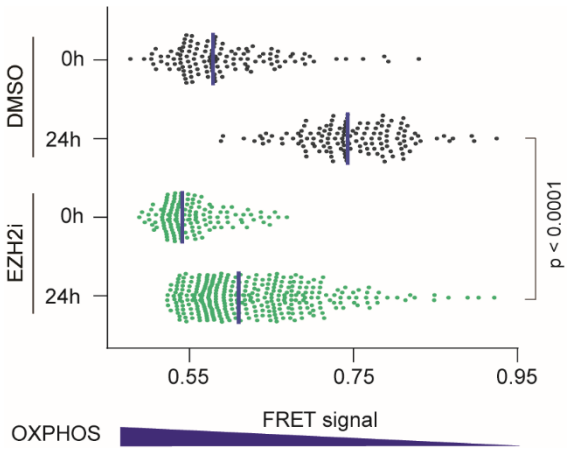


Figure 42: Epigenetically disrupted cells resist to the pressure posed by nutrient starvation

Quantification of oxidative phosphorylation (OXPHOS) levels in living MEXF 2090 cells grown for 24h under nutrient starvation using the FRET-based biosensor. 126 < N < 330 cells. P-value from two-tailed Mann–Whitney U-test.

Aiming to directly assess cell state transitions at the single cell level, I manually tracked the metabolic state of individual cells via time-lapse imaging in 15-minute intervals during their early response to starvation (24h). It is worth noting that this time frame represents the initial stress response and precedes any detectable phenotypic alterations like halting of

proliferation and cell death. Over the first 24 hours post exposure to starvation, DMSO treated cells gradually transitioned towards a stressed OXPHOS^{low} state, while EZH2i treated cells at some point diverged and retained higher mitochondrial activity [Figure 43A-B]. The detection of overall similar patterns between the single cell and population analysis further solidifies these observations as it suggests that the sampling of a limited number of individual cells did not bias the observed cellular behaviors.

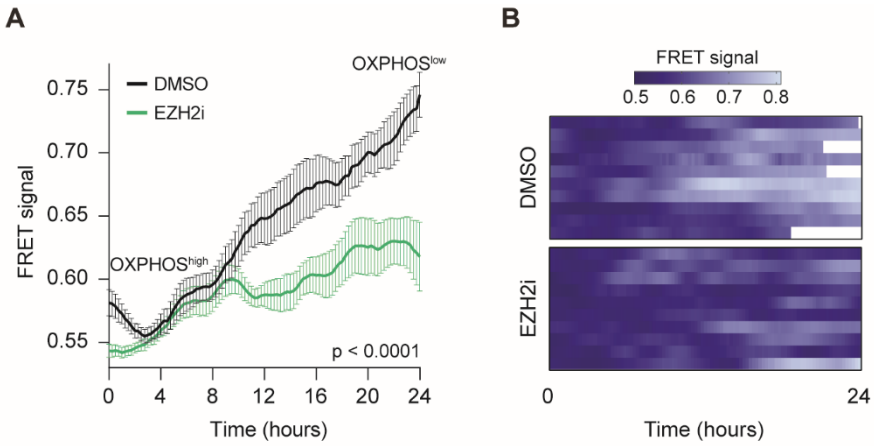


Figure 43: Unaltered plasticity of single cells upon disruption of the epigenetic network

[A-B] Time-lapse imaging of the indicated cells. Values in A represent mean ± SEM from 10 cells, shown separately in the heatmap in B. P-value from two-tailed Student's t-test calculated at the last time point. White squares in B represent time points where cells could not be tracked.

Overall, the data obtained by monitoring of metabolic state transitions in response to nutrient starvation provide initial evidence against the model of enhanced phenotypic plasticity in epigenetically disrupted cancer cells. On the contrary, EZH2i cells were less responsive to their changing environment.

4.3.3 Phenotypic inertia of epigenetically disrupted cells

Note: The single-cell RNA sequencing experiment was performed by the Advanced Sequencing Facility at the Francis Crick Institute. Paolo Inglese performed a) pre-processing of the raw data, and b) cell clustering and identification of KO-enriched or control-enriched subpopulations. The subsequent analysis was performed by Ioannis Loukas and Paola Scaffidi.

4.3.3.1 Multiplexed single-cell RNA-seq in unperturbed conditions and under nutrient starvation

The data presented so far suggest that genomic instability cannot explain the stress resistance of epigenetically disrupted cells. On top of that, initial evidence from live-cell imaging along with transcriptomic analysis at the population level argue against the phenotypic plasticity or bet-hedging model. Aiming to further explore aspects of the system in greater detail and gather data in favor of the potential remaining models I decided to perform single-cell transcriptomic analysis (scRNA-seq), comparing the alterations induced by nutrient starvation in control and epigenetically deregulated cells at distinct time points.

To minimize the inherent technical noise of scRNA-seq experiments that can lead to significant batch effects, I decided to perform multiplexed transcriptomic analysis via the 10X Genomics' Feature barcoding technology. This approach allows for the simultaneous interrogation of the transcriptome of multiple populations that are subsequently demultiplexed based on the distinct sgRNA transcripts that they express. Such resolution is achieved via the presence of two different classes of oligonucleotides on the 10X Genomics gel beads that can simultaneously capture both the total polyadenylated mRNA and the sgRNA transcripts from individual cells [Figure 44A]. In greater detail, the profiling of the sgRNAs is mediated through the hybridization of specialized oligonucleotides on the surface of the beads that carry a capture sequence that is complementary to the sgRNA scaffold [Figure 44B]. The scaffold within the sgRNA library that I have used so far does not carry this sequence (Henser-Brownhill, Monserrat & Scaffidi, 2017) and thus it is not compatible with such a single-cell transcriptomic pipeline. To generate new sgRNA scaffolds, I used synthetic constructs and cloned them into the pLenti-sgRNA plasmid that drives the expression of the sgRNAs once inserted into the cells (For more details see Methods section 2.3.2). The successful cloning of the engineered scaffold was confirmed by sanger sequencing [Figure 44C]. Finally, I selected the KO populations used earlier for the bulk RNA-seq and after lentiviral transduction with the engineered scaffolds, I generated sgRNA barcoded populations compatible with the subsequent experimental pipeline.

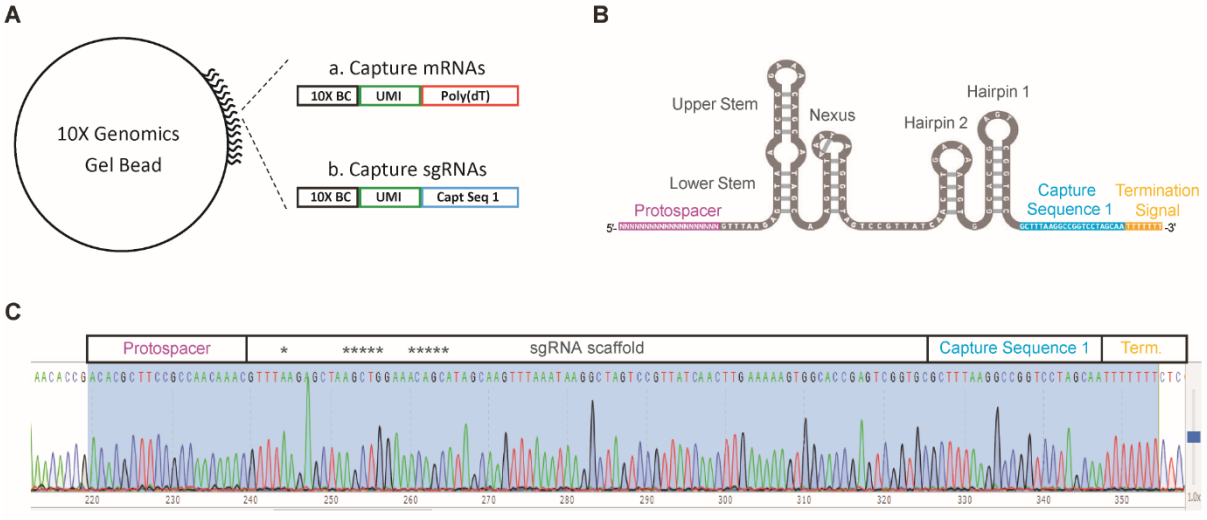


Figure 44: Cloning of a modified sgRNA scaffold that is compatible with 10X Genomics' Feature barcoding technology

[A] Schematic representation of a 10X Genomics gel bead. Note the presence of two distinct classes of oligos which provide the capacity to capture per cell both poly(A) transcripts and sgRNAs. UMI: Unique molecular identifier.
[B] Schematic representation of the sgRNA scaffold carrying the Capture sequence 1 at the 3'end of the sequence, immediately before the termination signal.
[C] Sanger sequencing confirming the successful generation of the engineered sgRNA scaffold carrying the Capture sequence 1.

The transcriptomic analysis was performed at 4 different time points:

- Day 0 [do]: before the application of stress (unperturbed condition)
- Day 1 [d1]: early response to starvation before the detection of any phenotypic alterations on proliferation or viability induced by nutrient deprivation [Figure 45A].
- Day 2 [d2]: early response to starvation where the first indication of impaired growth is visible, but no significant effect on cell viability [Figure 45A].

  The absence of cell death in the first two time points, suggests that any detected alterations are the result of transcriptomic rewiring rather than selection.
- Day 12 [d12]: Late response to starvation and selection of resistant or adapted subpopulations. At this time point, all the epigenetically disrupted cells demonstrate increased fitness at varying degrees compared to the control cells [Figure 45B].

To further minimize the potential technical noise, the experimental set-up included a reverse time-course, where all the samples to be analyzed were sequenced at the same time [Figure 45C].

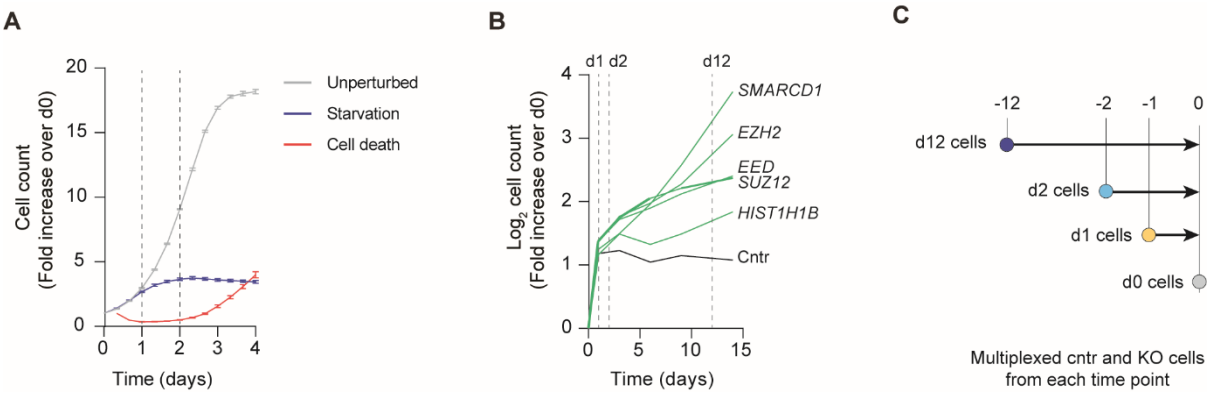


Figure 45: Multiplexed scRNA-seq via sgRNA barcoded populations in unperturbed and nutrient starved conditions

[A] Growth kinetics of MEXF 2090 cells stably expressing GFP in unperturbed conditions (grey) or under nutrient starvation (blue). Red line indicates the relative cell death observed in nutrient starvation, as assessed by monitoring apoptosis via a Caspase-3+ fluorescent indicator. Values represent mean ± SEM from three biological replicates.
[B] Growth kinetics of the KO populations of MEXF 2090 cells analyzed by scRNA-seq under nutrient starvation. The time points selected for analysis, in addition to unperturbed cells at d0, are indicated.
[C] Schematic representation of the experimental design (reverse-time course) followed during the single-cell transcriptomics analysis of MEXF 2090 cells grown under nutrient starvation for the indicated number of days.

Considering that one of the main objectives of the subsequent analysis was to assess transcriptional heterogeneity within the populations (for the plasticity and bet-hedging models), I went on to sequence the samples at a significant depth, reaching ~87.000 reads/cell, which resulted in the detection of ~5800 genes/cell (The above values represent mean estimates across all samples and time points). For each KO population at each time point, the profiled cells ranged between 270 and 670. The generated datasets from single-cell transcriptomics experiments often contain “low-quality” cells that are the by-product of the stochastic Poison sampling from the population along with the capturing of stressed, broken, or dead cells. These cells can be identified through a combination of criteria such as the low number of detected transcripts and enriched mapping of sequencing reads to mitochondrial-DNA genes (See also Methods section 2.16.3.1). Quantification of such low quality / stressed cells (from here on referred to as LQ cells) failed to detect any changes in the number during the immediate response to starvation (d1 and d2), in line with previous observations where no differences in viability were detected during the first 48 hours in starvation [Figures 45A and 46]. On the contrary, at d12 there was a significant increase in stressed LQ cells, with more than 20% of the low-quality cells belonging to control MEXF 2090 cells. Epigenetically disrupted populations maintained lower fractions of stressed cells, in correlation with the extent of survival advantage observed in the fitness assays [Figures 45B and 46]. Low quality cells can hinder the downstream analysis in various ways, thus they were discarded from the dataset. Consequently, the detected differences among the control and the KO populations

are an underestimation of the actual effects, especially in the case of *HIST1H1B*-deficient cells that exhibit only a mild fitness advantage compared to the control population. Overall, the above data suggest that the generated dataset is of good quality, faithfully recapitulates the detected differences from the fitness assays and is worth further investigation.

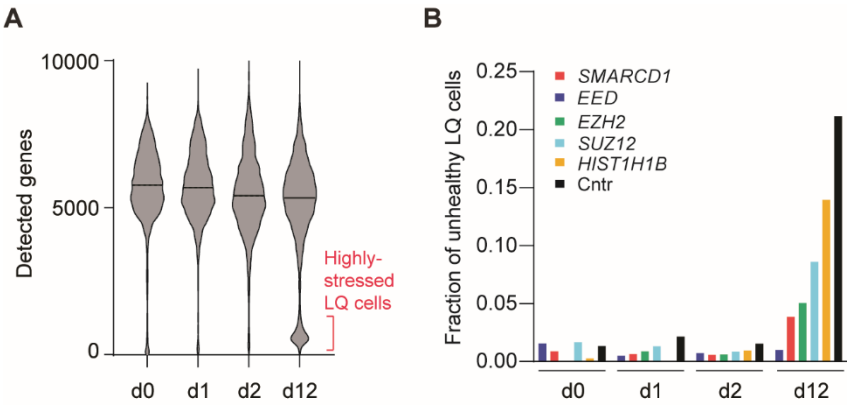


Figure 46: Detection of low quality (LQ) stressed cells per time point and KO population

[A] Quantification of genes detected in all cell populations (merged samples) at the indicated times, showing an enrichment of highly stressed, low-quality (LQ) cells at d12. 3112 < N < 4799 cells.
[B] Quantification of LQ cells in the indicated KO populations. 518 < N < 799 cells.

4.3.3.2 Dissecting the transcriptomic response to starvation at different time points

I first explored the stress response of control MEXF 2090 cells to nutrient starvation. UMAP based dimensionality reduction revealed that cells from different time points grouped in distinct clusters organized in a stepwise manner along the time trajectory [Figure 47]. This clear separation among the clusters suggests that the selected time points can robustly infer the distinct cellular behaviors of cancer cells that face starvation, from sensing and early response to stress to the subsequent alterations in cellular behavior that mediate the survival and propagation under chronic nutrient starvation.

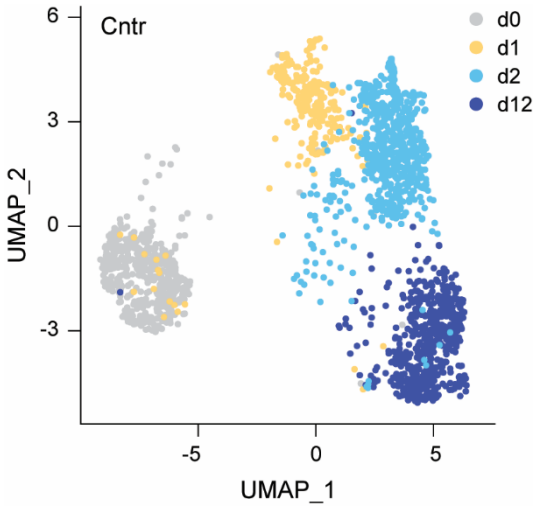


Figure 47: UMAP illustrating the changes induced by nutrient starvation in control MEXF 2090 cells

Assessment of the genes that get differentially regulated upon stress showed a gradual increase in the number of DEGs, in line with the above observations [Figure 48A]. To better understand the nature of such alterations I selected the 2340 genes that get differentially expressed during the first 48 hours in nutrient starvation (d0 vs d2) and explored their nature through GSEA analysis. In line with the results from the bulk RNA-seq experiment, the early response to starvation involved a significant downregulation of fitness related signatures (cMYC targets, oxidative phosphorylation and cell cycle genes) and the simultaneous upregulation of stress related pathways (p53 and NFkB target genes) [Figure 48B]. At the latest time point (DEGs between d2 and d12), a strong inflammatory response was observed along with additional upregulation of stress signatures (e.g. apoptosis), but also a partial reversal of transcriptomic changes that were detected at d2. The *reversed* genes include the re-expression of mTOR targets along with the attenuation of the p53 response [Figure 48C]. The latter is in line with previous observations that p53 mediated signalling is important for the early but not the late response to nutrient starvation (Tajan et al., 2018). The reversal of some early transcriptomic changes induced by the stress is a hint towards the enrichment of cells with the capacity to combat stress either as a result of *de novo* adaptation/rewiring or through selection of pre-existing stress-resistant subpopulations.

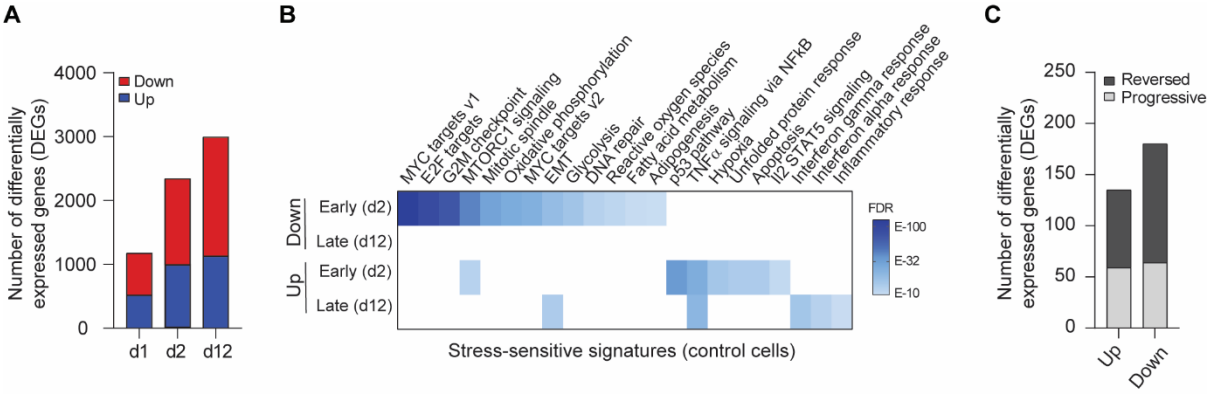


Figure 48: Dissecting the early and late transcriptomic changes in response to nutrient starvation

[A] Quantification of upregulated (up) or downregulated (down) genes in control cells at the indicated times after stress. DEGs are defined based on comparisons with cells at d0.
[B] GSEA analysis of early (d2 vs d0) and late (d12 vs d2) DEGs detected in control cells. The heatmap visualizes the significance of the enrichment of the indicated signatures. Signatures with FDR ≤ 10^{-10} are considered enriched.
[C] Quantification of genes showing progressive upregulation or downregulation from d0 to d2 and from d2 to d12, or reversal of the trend between d2 and d12.

To better understand the observed heterogeneity within the control populations in response to starvation, I calculated the pathway scores (i.e. weighted estimate of the expression of a specific gene set in individual cells) from the previously identified stress responsive signatures. This analysis revealed a) extensive cell to cell heterogeneity in both stress and fitness pathways b) a combination of both progressive and partially reversing trends and c) signs of resumed proliferation in a fraction of cells, as indicated by a higher median value of cMYC signature score at d12 [Figure 49]. These observations are in line with the results from the fitness assays and provide molecular evidence that even at d12 the majority of the control population succumb to stress but a few stress-resistant, proliferating cells are selected over time [Figures 45B, 48C and 49].

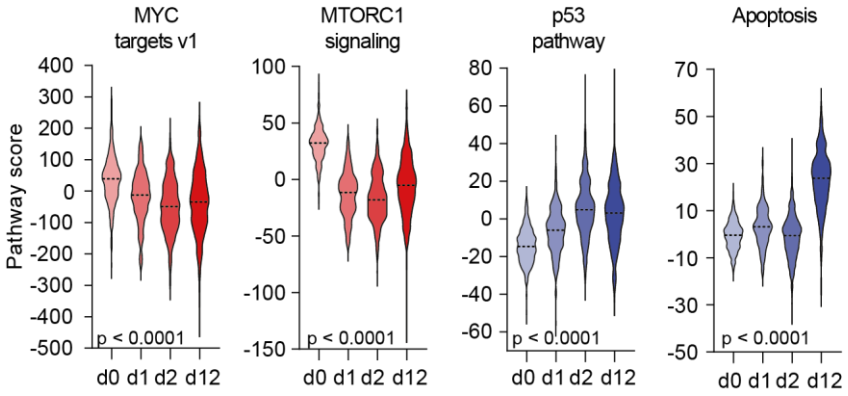


Figure 49: Pathway scores of fitness and stress signatures

Distribution of pathway scores in individual cells for the indicated gene signatures in control cells at the indicated times. P-values from one-way ANOVA assessing overall difference over time. 270 < N < 653 cells.

4.3.3.3 Absence of shared adaptive signatures in stress-resistant epigenetically disrupted cells

Inspection of the trajectories followed by the epigenetically disrupted populations in response to starvation, as defined through UMAP dimensionality reduction, revealed comparable patterns with the control cells [Figure 50]. Thus, most of the cells within the epigenetically disrupted populations also suffer under stress. Visually, the only significant difference between control and KO populations was detected at d0 (unperturbed conditions) where PRC2 cells cluster separately due to the de-repression of hundreds of developmental genes [Figures 34 and 50]. These genes constitute canonical PRC2 targets and have minimal overlap (1.9%) with the stress responsive genes, suggesting that they are irrelevant to the phenotypes observed during nutrient starvation. The above observations agree with the bulk RNA-seq analysis and reinforce the notion that stress-induced gene expression changes are dominant over regulator-specific effects induced by KO of specific genes [Figure 36].

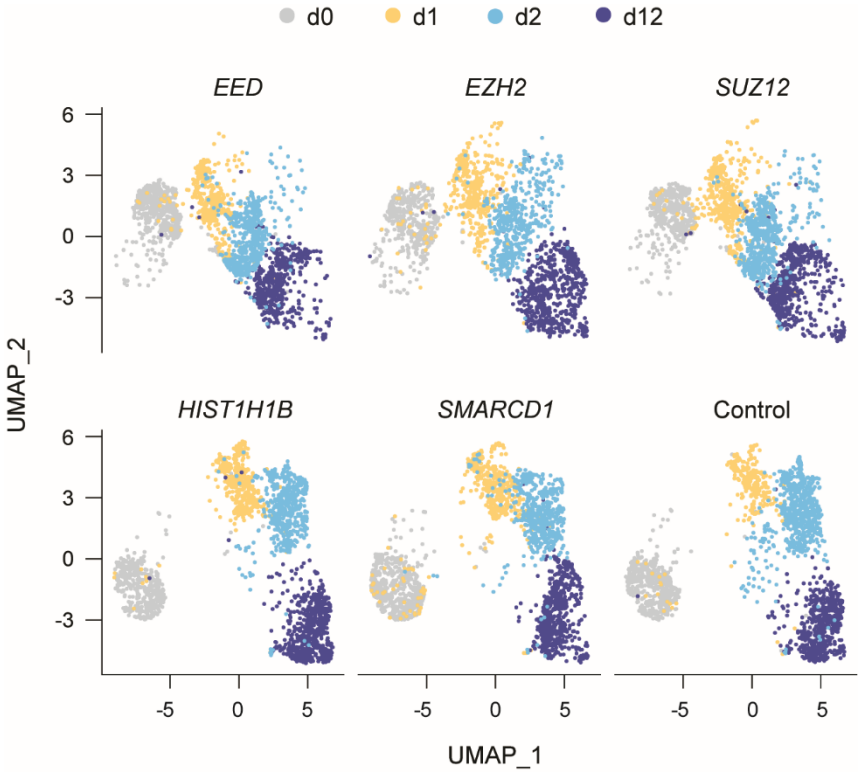


Figure 50: UMAP illustrating the trajectories of KO populations in response to nutrient starvation

Having established that all populations after prolonged growth in nutrient starvation (d12) contain a mixture of stressed and resistant cells in varying ratios [Figures 49 and 50], I went on to further dissect this heterogeneity. To do so, I focused on the latest time point (where we can observe the maximal phenotypic differences between the examined populations) and employed a clustering analysis looking for groups of cells that are either enriched for epigenetically disrupted cells (expected to consist mostly of resistant subpopulations) or for control cells (expected to contain stressed subpopulations). This approach involves pair-wise comparisons between control and epigenetically disrupted populations [Figure 51]. In brief, cells are mixed (irrespective of their KO status) and undergo unbiased clustering to identify groups of cells with distinct transcriptional behaviors. For each defined cluster a ratio of control over KO cells is calculated. The clusters that are differentially enriched for control or KO cells are the ones expected to drive the phenotypic differences observed in the fitness assays. Thus, for such clusters a differential gene expression analysis is performed to identify gene signatures specific to these groups of cells. After repeating this pipeline for all the relevant comparisons (each KO vs control), the extracted gene signatures are undergoing again unbiased clustering to explore the possibility that common meta-signatures are shared among different subpopulations of cells. The above analysis aimed to dissect the population heterogeneity and explore the presence of subsets of cells that display activation of common pathways that could provide cells with the capacity to propagate under stress. It is worth noting

that such adaptive signatures were not detected by the bulk RNA-seq analysis described earlier (see section 4.3.2.2).

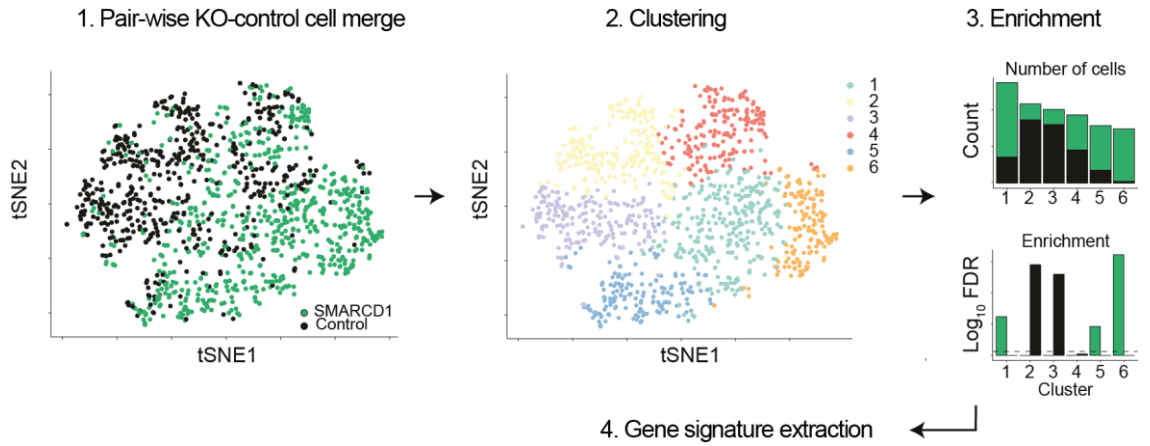


Figure 51: Dissecting population heterogeneity through meta-signature analysis

Unbiased clustering of the extracted signatures revealed the presence of two major meta-clusters. Each of them is predominantly occupied by either KO-enriched or control-enriched subpopulations [Figure 52A]. This dichotomous segregation suggests that distinct KO populations contain subsets of cells that are characterized by common transcriptional traits that are not shared by control cells. However, I did not detect the expression of any gene signature that could promote the adaptation of cancer cells to nutrient starvation, such as reprogramming to alternative differentiation states with differential sensitivity to glutamine deprivation or upregulation of other metabolic pathways that could act in a compensatory way or alternative mechanisms irrespective of metabolic rewiring like the activation of micropinocytosis (Zhang, Pavlova & Thompson, 2017). It is worth noting that the above analysis cannot exclude the possibility that the fitness advantage of the KO populations is conferred by distinct small sets of genes that are different in every single population.

To better understand the nature of the emerging meta-clusters, I curated gene sets that are highly correlated with either of them and interrogated their nature through GSEA analysis. The KO-enriched subpopulations were characterized by elevated expression of fitness signatures and simultaneously decreased levels of stress associated pathways, suggesting that they represent stress-resistant cells. Conversely, the control-enriched meta-cluster was characterized by an antithetic pattern of gene expression [Figure 52B]. The above data strengthen the notion that after prolonged culture in nutrient starvation epigenetically disrupted cells retain a greater fraction of stress-resistant subpopulation.

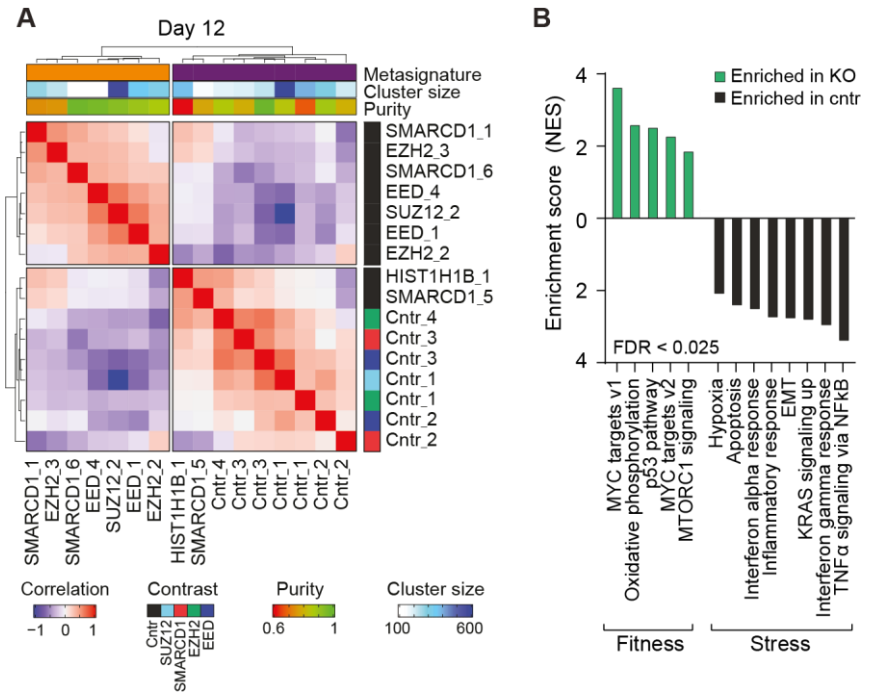


Figure 52: KO-enriched and control-enriched subpopulations occupy distinct transcriptional states

[A] Unbiased clustering of gene signatures defining the indicated KO- or control-enriched subpopulations at d12. Multiple enriched subpopulations identified in each pairwise comparison are indicated by progressive numbers.
[B] GSEA analysis of meta-signatures identified in A, showing a more fit and less stressed phenotype of epigenetically disrupted cells.

There are two distinct scenarios that could explain such behavior. One possibility is that epigenetically disrupted cells elicit a full stress response and then they manage (through secondary mechanisms) to revert to a less stressed state (in line potentially with a more phenotypically plastic population). Alternatively, they could be characterized by pre-existing increased tolerance to stress, enabling them to never reach a stress maximum as defined in the control cells, thus lowering the probability of cell death. Following the behavior of representative stress (*IL1B*) or fitness (*RPS26*) related genes over the time trajectory revealed that the latter is the case [Figure 53A]. Accordingly, a similar pattern was observed when the fitness signatures OXPHOS and Myc targets were examined along the time trajectory [Figure 53B].

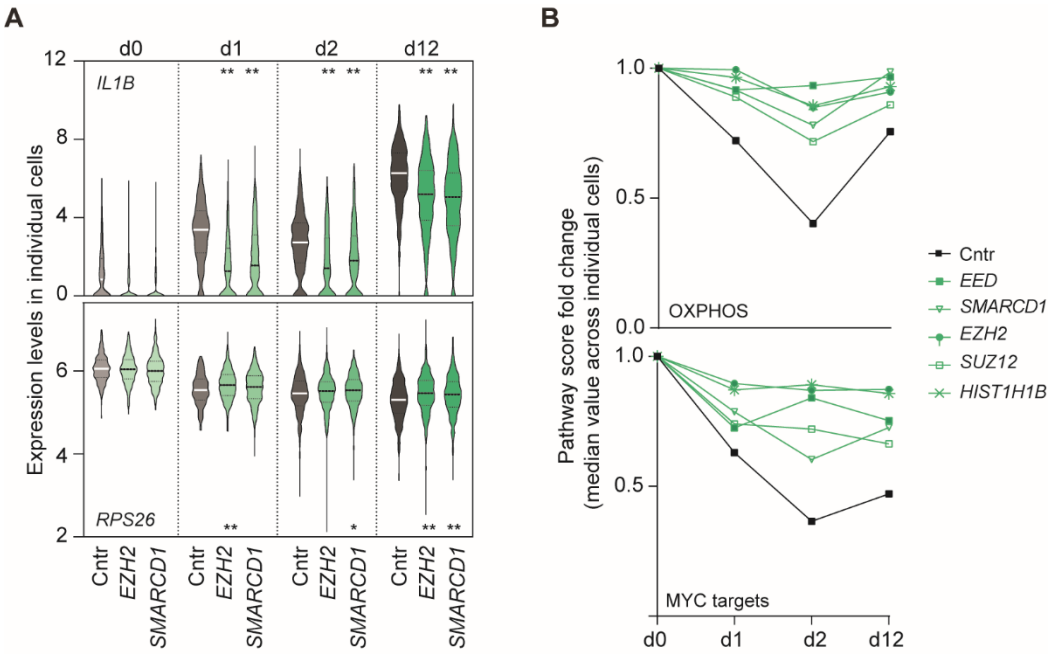


Figure 53: Trajectories of stress responsive genes / pathways in control and epigenetically disrupted cells

[A] Expression levels of a stress gene (*IL1B*) and a Myc-target (*RPS26*) in individual cells in the indicated populations, at the indicated times after nutrient starvation. P-values from Kolmogorov-Smirnov test. One and two asterisks indicate $p < 0.05$ and $p < 0.01$, respectively. $270 < N < 654$ cells. [B] Quantification of the relative changes in the indicated pathway scores over time upon stress in the indicated cell populations. The median value is plotted for each cell population.

Overall, dissecting the population heterogeneity of control and KO populations failed to detect the presence of adaptive signatures that can promote survival under nutrient starvation. This is in line with the observations from the live-cell imaging experiments and altogether indicate that the stress dependent advantage of epigenetically deregulated cells is not a result of a more efficient rewiring of cellular states. On the contrary, KO cells seem to be less responsive to the alterations induced by stress, consistent with the possibility that pre-existing stress-tolerant cells may be selected over time. Altogether the data presented so far argue against the model of increased plasticity or transcriptional variability of epigenetically disrupted cells and provide evidence that they are more inert when exposed to stress.

4.3.4 Diversification bet-hedging is not enhanced in epigenetically disrupted cells

Note: Phil East applied the SCDE/PAGODA algorithm to the data in order to estimate the transcriptional variance per gene. The subsequent analysis was performed by Ioannis Loukas and Paola Scaffidi.

4.3.4.1 Robust inference of highly-variable genes from scRNA-seq data

One of the most interesting aspects of single-cell transcriptomics is that it allows the dissection of the transcriptomic variance within a seemingly homogenous population, an important aspect of many biological contexts that is lost during bulk analysis. The presence of such cell-to-cell variability in gene expression increases the probability that a subset of cells expresses advantageous genes that leads to the acquisition of a favorable phenotype in specific contexts. This variability is the molecular underpinning of bet-hedging, a strategy employed by various organisms to promote phenotypic diversity and thus maximize their potential to survive in unpredictable environments. Conceptually, in the context of cancer bet-hedging can be considered the transcriptomic equivalent of genomic instability, which leads to intratumor phenotypic diversity upon which selection can act. Such transcriptomic variability has been previously implicated as an important driver of resistance to stress (Shaffer et al., 2017).

So far, the analysis presented in this chapter has failed to detect evidence of such a transcriptomic state, that emerges upon epigenetic deregulation and is strongly favorable during the stress response. It is worth noting though, that this analysis has primarily focused on the latest time point (day 12), based on the assumption that after chronic exposure to stress this rare favorable transcriptional state would be selected, enriched and thus detected by single-cell transcriptomics. To directly assess if epigenetic deregulation affects the transcriptional variance of cess, I next focused on the early time points and specifically before the application of stress (day 0). In greater detail, I compared the gene expression variability, as quantified by the calculation of the coefficient of variation (CV^2) for each gene, between control and epigenetically disrupted cells in unperturbed conditions. Genes that are lowly expressed (noisy) are characterized by increased coefficient of variation while highly expressed genes exhibit the opposite behavior. To correct this dependency, I utilized the SCDE / PAGODA algorithm that can normalize on a gene-by-gene basis the observed variance to the expected one based on the genome-wide properties. On top of that, PAGODA corrects for additional factors that can hinder the faithful estimation of expression variance like gene length (Fan et al., 2016; Faure, Schmiedel & Lehner, 2017).

Initially, I focused on control cells growing in unperturbed conditions and curated a list of 428 highly-variable genes (HVGs), by selecting the 95th percentile of normalized variance as a cut-off (genes with top 5% variance)[8]. These genes were characterized by relatively high mean

[8] The rationale behind this approach was to mathematically define in control cells a threshold that characterises genes with large variance in their expression levels and then use this numerical value as a cut-off to explore potential differences (quantitative and qualitative; number and nature of HVGs respectively) across different

expression and were detected in the majority of the cells, suggesting that the defined genes are the ones with significant expression variability within the population rather than rare noisy expression [Figure 54A-B]. GSEA analysis revealed that HVGs are significantly enriched for stress signatures (NFkB targets, Apoptosis, etc) and genes involved in cellular plasticity (EMT), in line with the idea that priming subsets of cells for prompt reaction to unfavorable environments enables an efficient response at the population level [Figure 54C]. On the contrary, fitness pathways were characterized by even higher mean expression but minimal variability. This segregation among stress and fitness signatures reflects an underlying selective pressure to evolve common regulatory traits ensuring that functionally related gene sets behave in similar ways (expression level, variability, etc.) [Figure 54D].

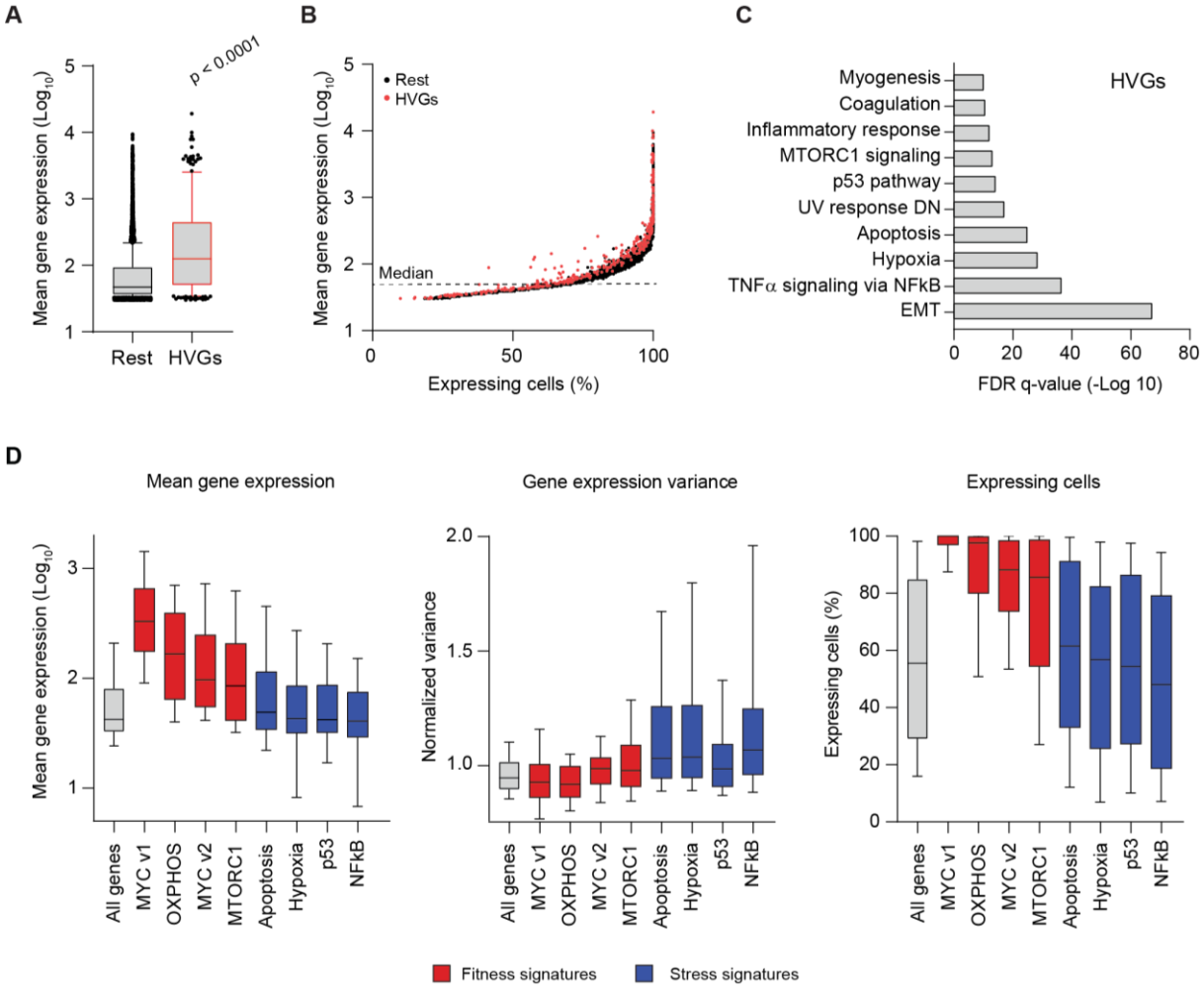


Figure 54: Characteristics of highly-variable genes (HVGs)

[A] Distribution of mean expression levels of HVGs compared to all other genes (rest) in control unperturbed cells. The bottom and top of boxes indicate the 25th and 75th percentiles, respectively,

populations (e.g., control vs epigenetically deregulated cells). Other metrics of variability (cell-based) were also assessed that are not presented in this thesis (for more information refer to Discussion section 5.4).

and middle lines indicate medians. Whiskers indicate the 5th and 95th percentiles. Pagoda-normalized values are used. P-value from two-tailed Mann-Whitney U-test. N = 428 (HGVs), 8133 (rest) genes.
[B] Relationship between mean expression levels and percentage of cells expressing a given gene, with HGVs labelled in red.
[C] GSEA analysis of HVGs. Signature with FDR ≤ 10^{-10} are considered enriched.
[D] Features of the indicated gene signatures showing distinct patterns for stress and fitness signatures. The bottom and top of boxes indicate the 25th and 75th percentiles, respectively, and middle lines indicate medians. Whiskers indicate the 10th and 90th percentiles, with outliers omitted for clarity.

4.3.4.2 Epigenetic deregulation does not affect transcriptional variance

Having confirmed the robust inference of highly-variable genes from the single-cell transcriptomics dataset, I then asked if epigenetic deregulation could alter the properties of HVGs and more specifically look for indications of greater phenotypic diversity in the absence of stress. Thus, by using the variance cut-off defined in the control cells in unperturbed conditions I quantified HVGs in the KO populations. No significant difference was observed in the number of HVGs in epigenetically deregulated cells upon epigenetic deregulation [Figure 55A-B]. When quantifying the extent of the variance of HVGs I observed a slight deviation towards higher variance in two of the three components of the PRC2 complex [Figure 55C]. In most of the analysis that I have presented so far *EDD*- and *SUZ12*-KO populations behave in an identical manner. Thus, the observed inconsistency challenges the biological importance of the observed difference. Considering that the number and the variance are not altered, I then asked if the nature of these genes is different among the populations. I observed a significant overlap where more than 40% of the HVGs were shared among control and KO cells [Figure 55D]. Notably, GSEA analysis of these genes revealed the enrichment of the same signatures, suggesting that the KO populations express in a variable manner common but also distinct components of the same pathways (data not shown).

While I cannot rule out the possibility that specific genes relevant for survival under nutrient starvation acquire higher variability in a KO specific manner, it is unlikely that the increased phenotypic diversification in unperturbed conditions via bet-hedging is the underlying mechanism mediating the selective advantage of epigenetically disrupted cells under unfavorable conditions.

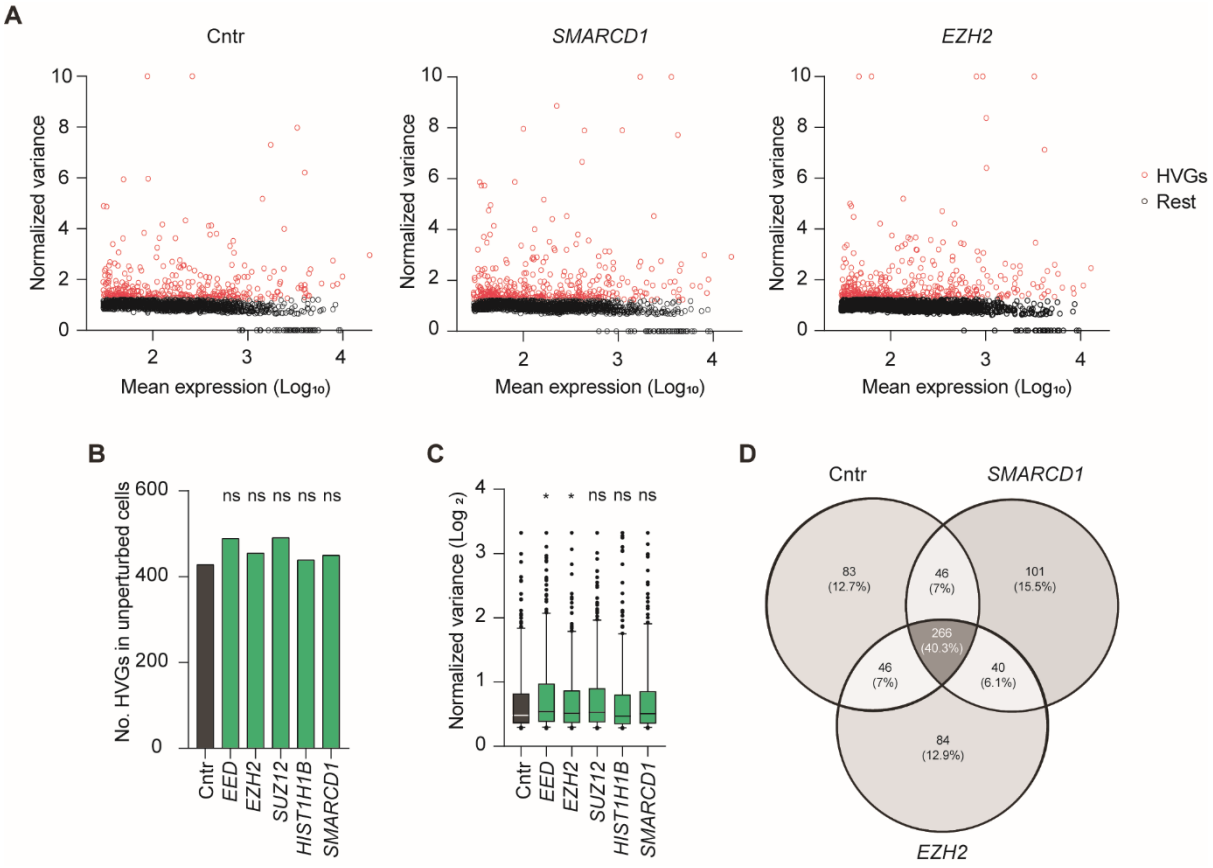


Figure 55: Unaltered transcriptional variance in epigenetically disrupted cells under unperturbed conditions

[A] Normalized variance of gene expression levels in the indicated populations. Each dot is a gene, with HVGs indicated in red.
[B-C] Quantification of the number of HVGs (B) and their normalized variance (C) in the indicated samples in unperturbed conditions. P-value relative to control cells from two-tailed Fisher test (B) or Mann-Whitney U-test (C).
[D] Venn diagram showing the overlap of HVGs in the indicated populations.

4.3.5 Chromatin-mediated changes in global transcriptional activity in response to stress

Note: Paolo Inglese run the txburst algorithm to estimate the bursting properties. The subsequent analysis was performed by Ioannis Loukas and Paola Scaffidi.

4.3.5.1 Inference of transcriptional burst properties from scRNA-seq data

How is it possible that mutations in epigenetic regulators with distinct functional and structural roles within the chromatin can result in the same outcome, phenotypic inertia? The analysis presented so far has revealed that nutrient starvation induces significant alterations in the

transcriptional patterns of melanoma control cells, involving changes in functionally related fitness and stress signatures. Epigenetically disrupted populations follow the same pattern but exhibit a milder phenotype. One possibility is that the KO populations cannot elicit a strong stress response at the transcriptional level, thus preventing them from halting proliferation and committing to apoptosis and subsequently cell death. Considering that the fitness advantage is observed in populations KO for distinct chromatin regulators, it is doubtful that the misregulation of specific gene sets mediates their behavior. Thus, I explored more universal features of transcriptional activity that could be broadly affected by the disruption of diverse members of the epigenetic network.

Transcription is not continuous but rather happens in episodic bursts (Levine, Lin & Elowitz, 2013). The stochasticity of the phenomenon is a result of the spontaneous interaction of the participating factors that are required to initiate transcription along with the limiting nature of the substrate (gene locus to be transcribed). While seemingly random, biological systems have developed mechanisms to tune its kinetics and thus alter the probability of the reaction taking place. These mechanisms involve fixed traits that emerged during evolution, such as sequence encoded elements within the promoter region. On top of that, the packaging of the DNA into chromatin can alter the accessibility to the substrate and or the effective concentration of the participating regulators thus providing another regulatory layer to transcription (Rodriguez & Larson, 2020; Tunnacliffe & Chubb, 2020).

The above fixed and dynamic regulatory traits result in a complex setting where every single gene can have a varying probability to fire. Each firing event is characterized by its frequency (i.e. how often a gene is transcribed) and its size (i.e. the amount of the mRNA produced in each burst). These parameters can be estimated from static scRNA-seq data, assuming a two-state model of transcription where promoter elements shuttle between an active and an inactive state at gene specific rates (Burst frequency: K_{on}, in mRNA degradation units; burst size: K_{syn}/K_{off}) [Figure 56] (Larsson et al., 2019).

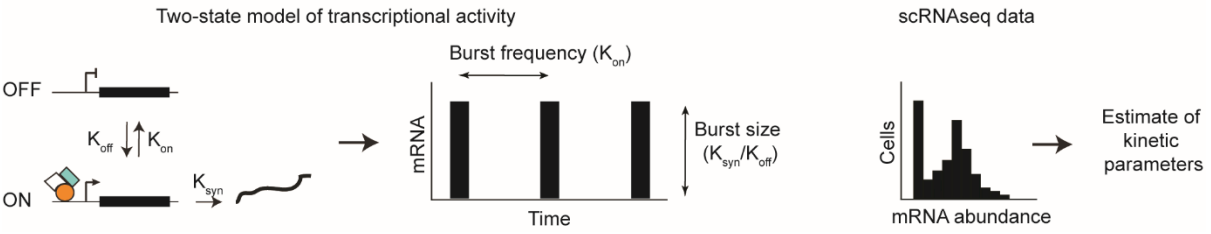


Figure 56: A two-state model of transcriptional activity used to infer transcriptional burst parameters from static scRNA-seq data

I went on to employ an approach that uses profile likelihood to estimate the transcriptional burst parameters (Larsson et al., 2019). Of note, there are some limitations associated with

applying this pipeline in my experimental dataset. Firstly, the approach described by Larsson et al. has been developed to infer burst kinetics from scRNA-seq experiments with single allele resolution, not in line with the nature of my dataset. After analyzing the data from Larsson et al, I observed a significant correlation between the quantifications obtained from the different alleles of the same genes, suggesting that merging of the data can only over or underestimate the properties but not distort their relative ranking [Figure 8] (Larsson et al., 2019). Secondly, it is worth noting that in response to a strong environmental stimulus, a two-state model of transcription may fall short to capture all the complex variables at play (Tunnacliffe & Chubb, 2020). On top of that, stress per se can significantly affect the RNA turnover, a factor that is important for the accurate estimation of the bursting parameters. However, the analysis presented in this chapter is not meant to infer absolute estimates of bursting but rather explore relative alterations and use them as a proxy of global transcriptional activity across populations. Thus, acknowledging the potential limitations, I applied this algorithm to my dataset aiming to a) characterize changes in transcriptional burst properties induced by stress and b) ask if the disruption of epigenetic control can affect this response.

The relationship observed between the expression and the inferred burst parameters at the global scale under unperturbed conditions recapitulated what was previously reported by other groups, confirming the reliability of the generated dataset [Figure 57A] (Larsson et al., 2019; Ochiai et al., 2020). Notably, burst frequency is a stronger predictor of gene expression levels [Figure 57B-C]. This can be attributed to the presence of a dichotomous relationship between size and expression where a fraction of genes exhibits high burst sizes that do not lead to significant expression levels. GSEA analysis of these genes failed to detect any significant enrichment within known gene signatures (data not shown).

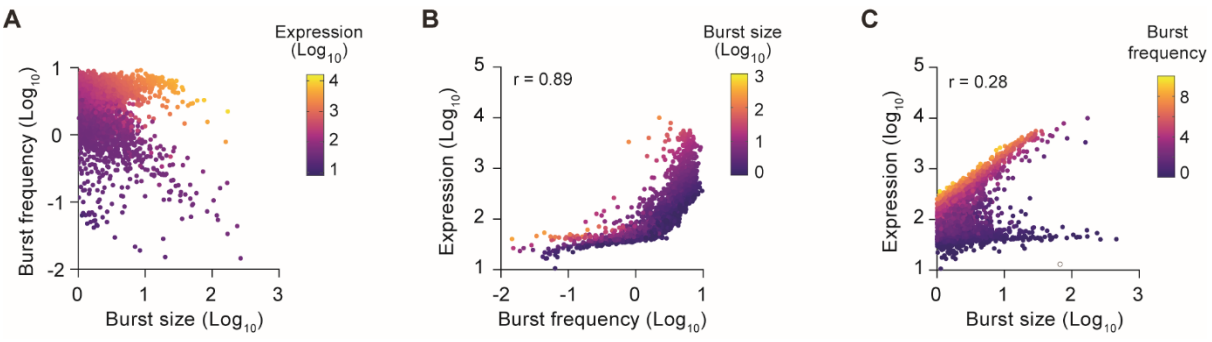


Figure 57: Relationship between gene expression and transcriptional burst parameters in unperturbed cells.

[A-C]Relationship between mean expression levels and transcriptional burst parameters in unperturbed control cells. Each dot is a gene. The Spearman correlation coefficient r is indicated.

4.3.5.2 Altered transcriptional burst properties induced by nutrient starvation

Having established confidence in the estimated burst properties, I then explored how nutrient starvation can affect the aforementioned relationships. Substantial alterations were detected in control cells after two days of growth in nutrient deprivation. Such stress induced global alterations in transcriptional activity, involved mainly a striking decrease in the transcriptional burst frequencies [Figure 58A]. Monitoring the frequency pattern over the time trajectory revealed a gradual decrease that reached a maximum after 48 hours under starvation followed by a partial reversal of the phenotype at the latest time point [Figure 58B]. The latter is in line with previous observations from fitness assays and the heterogeneity analysis of the scRNA-seq dataset indicating that control cells can rarely combat the effect of starvation leading to a less stressed state (potentially as a result of selection during the long-term exposure to stress). Pairwise comparison of the burst frequencies among the distinct time points revealed that the genes mainly affected are the ones characterized by high burst frequency [Figure 58D]. In contrast to the above changes, the distribution of burst size remained largely unaffected along the time trajectory except for 24 hours after stress exposure (day 1) where a minor increase was detected [Figure 58C].

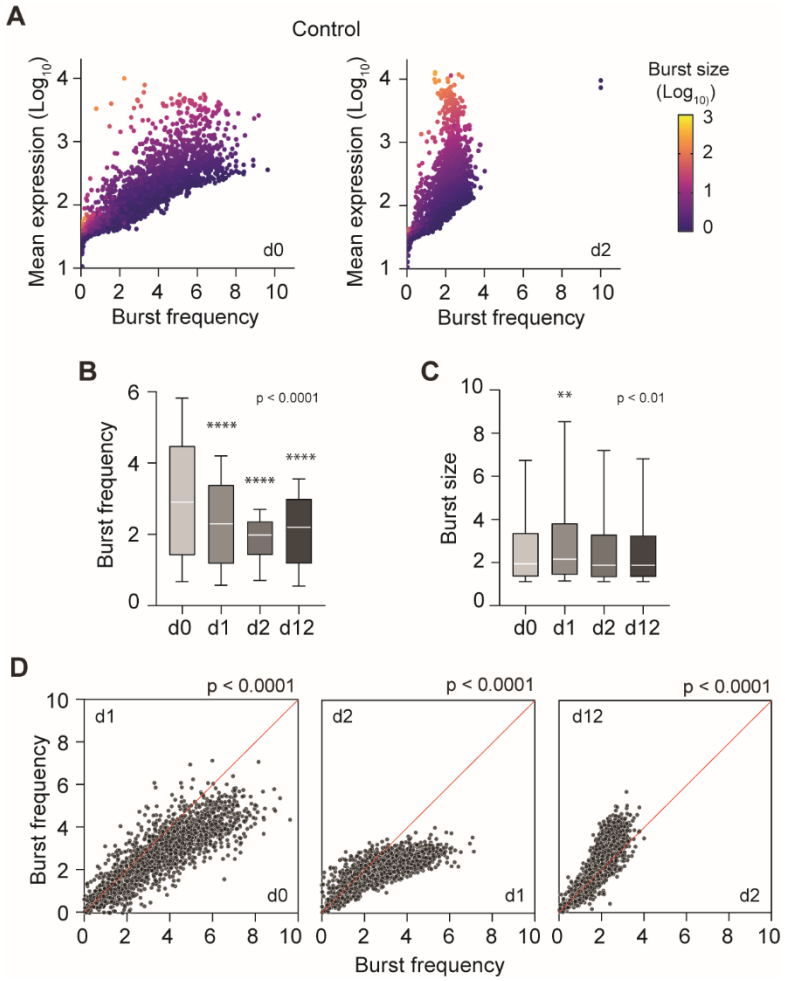


Figure 58: Global alterations in transcriptional burst properties upon nutrient starvation

[A] Visualization of the changes in transcriptional activity induced by nutrient starvation in control cells. Each dot is a gene.
[B-C] Distributions of transcriptional burst frequency [B] or size [C] for individual genes in control MEXF 2090 cells at the indicated time points under nutrient starvation. Whiskers indicate the 10th and 90th percentiles, with outliers omitted for clarity. P-value relative to d0 from one-way ANOVA followed by Dunnett's test for multiple comparisons (**p < 0.01 and ****p < 0.0001). 2814 < N < 3407 genes.
[D] Pair-wise comparison of transcriptional burst frequencies in control cells at the indicated times. Each dot is a gene. P-value from Wilcoxon test

To begin to understand how these alterations affect cellular behaviors I curated a list of high-frequency genes (HFGs, K_{on} > 4; top 33%) and explored their identity through GSEA analysis. Interestingly the HFGs were strongly enriched for Myc V1 target genes and components of the oxidative phosphorylation pathway and to a lesser extent for cell cycle related genes [Figure 59A]. The distribution of the burst frequencies of the above fitness signatures followed the same pattern, with a gradual decrease in response to stress being followed by a late reversal [Figure 59B]. Since the burst size remains unaffected, this plummeting in burst frequency of fitness related gene signatures provides a molecular framework that explains their downregulation over time in response to nutrient starvation. This phenomenon is likely a

pro-survival strategy for cells aiming at saving energy and biomolecules under unfavorable conditions (Pakos-Zebrucka et al., 2016).

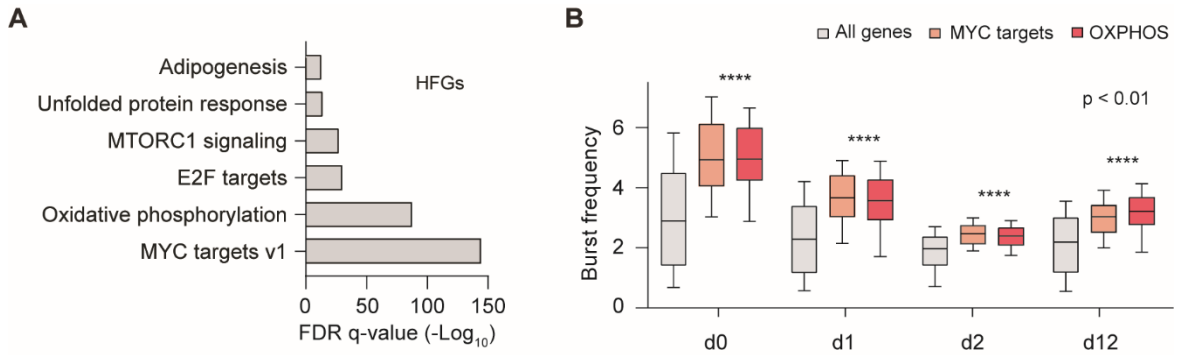


Figure 59: Reduction in burst frequencies primarily affects fitness signatures

[A] GSEA analysis of high-frequency genes, HFGs. Signatures with FDR ≤ 10^{-10} are considered enriched.
[B] Distributions of transcriptional burst frequency for the indicated gene sets in control cells. P-value from two-way ANOVA comparing overall trends of MYC targets or OXPHOS with All genes, followed by Sidak's multiple comparisons (****p < 0.0001). N ≥ 110 genes for MYC targets or OXPHOS.

Considering that nutrient deprivation induces a global decrease in transcriptional burst frequency, I wondered how stress signatures get upregulated over time. From the stress responsive genes (DEGs between d0 and d12), I curated a gene list by collapsing all the genes that get upregulated and belong to stress related pathways (e.g. Apoptosis, NFkB target genes, etc). Notably, these genes are characterized by low burst frequency in unperturbed conditions, making them less susceptible to the global frequency reduction observed upon stress and more reliant on burst size to modulate their expression [Figure 60A]. Indeed, the burst size of stress genes gradually increased over time reaching its maxima at the latest time point, in line with their upregulation [Figure 60B]. It is worth noting that this behavior is specific to the stress pathways and is in contrast with what is observed for the rest of the expressed loci [Figure 58B].

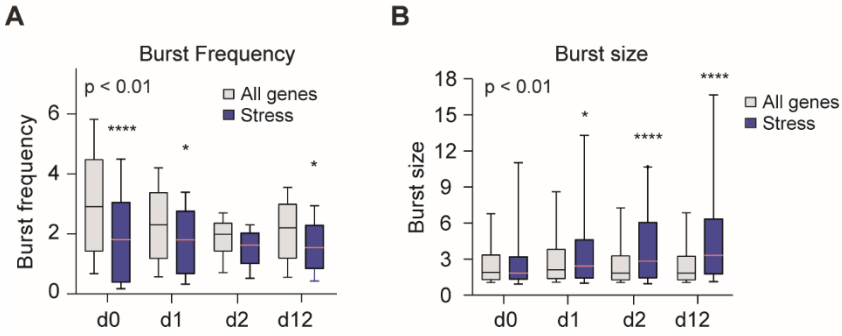


Figure 60: Selective increase in burst size of stress related genes

[A-B] Distributions of transcriptional burst frequency [A] and size [B] for all or stress genes in control cells at the indicated times under nutrient deprivation. P-value from two-way ANOVA, followed by Sidak's multiple comparisons (*p < 0.05 and ****p < 0.0001). N = 2814 (all) or 57 (stress) genes.

Altogether, the above data suggest that nutrient deprivation leads to substantial changes in transcriptional burst properties and more specifically to the downregulation of high-frequency fitness genes and simultaneously the upregulation of low-frequency stress genes through an increase in their respective burst size.

4.3.5.3 Epigenetically disrupted cells resist to the stress induced alterations in burst properties

Having established how nutrient starvation affects the bursting properties of fitness and stress related signatures, I then asked if disruption of the epigenetic network affects those trends. In contrast to the observed alterations in control cells, at day 2 all KO populations exhibited global patterns that were comparable to the ones before the application of the stress, suggesting a severely affected response to stress [Figure 61A]. Monitoring the frequencies over the time trajectory, revealed a gradual decrease upon stress application, reaching a maximum at day 2 and then showing signs of recovery. Although this trend is similar to the one observed in control cells, the extent of the alterations was significantly different in epigenetically disrupted cells that exhibited a much milder phenotype [Figure 61B-C]. The effect at d1 mirrored the differences observed in the fitness assays, with PRC2-defective cells diverging from control cells earlier than cells mutated for *SMARCD1* or *HIST1H1B* [Figure 45B]. Before the application of stress (day 0), KO populations showed comparable distributions of burst frequencies to the control cells, arguing against the possibility of pre-existing global differences in burst properties that make cells less susceptible to the induced global alterations thus priming them for a better response to stress [Figure 61D].

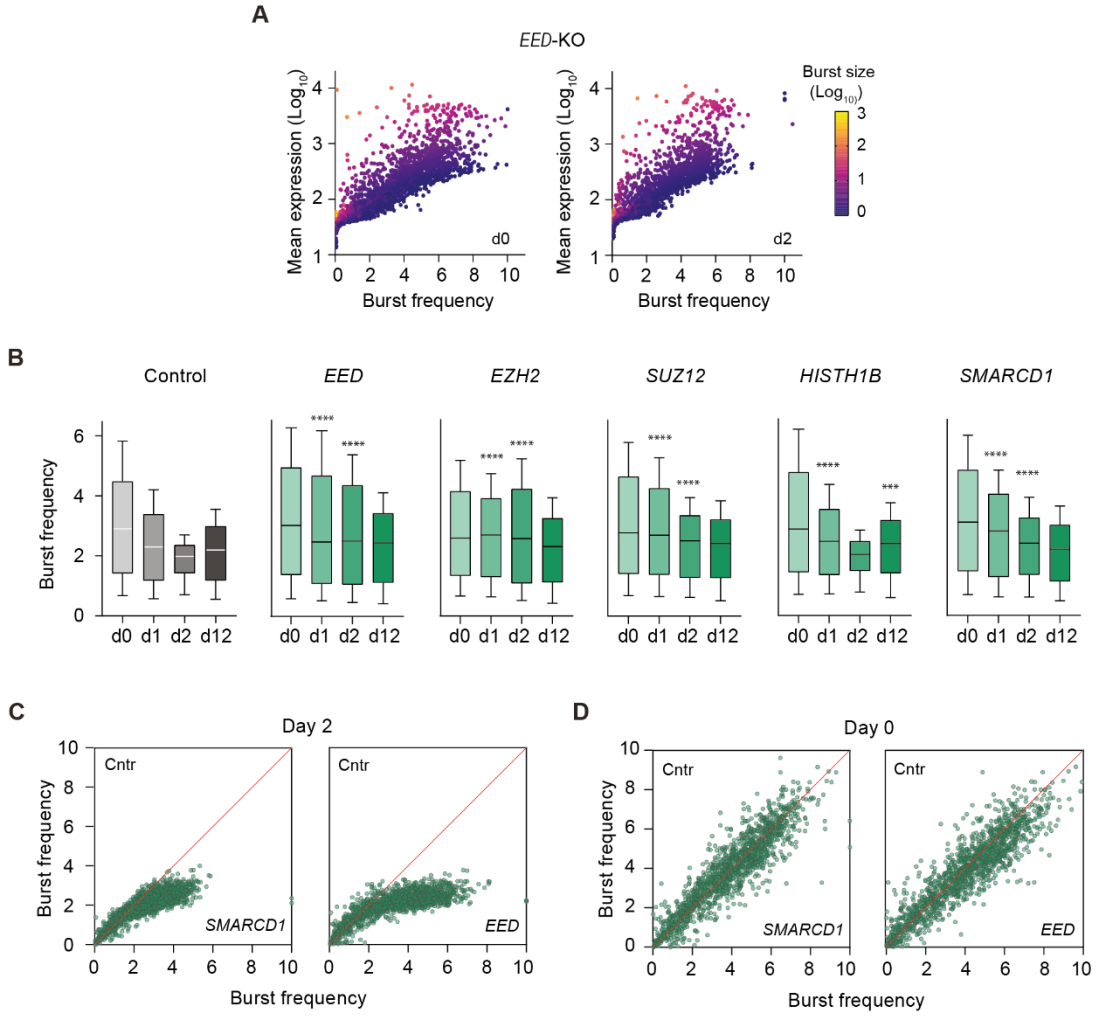


Figure 61: Epigenetically disrupted cells resist to the reduction of burst frequency

[A] Visualization of the changes in transcriptional activity induced by nutrient starvation in *EED*-KO cells. Each dot is a gene.
[B] Distributions of transcriptional burst frequency for individual genes. ***p < 0.001 and ****p < 0.0001 relative to the same time point in control cells (two-way ANOVA, followed by Sidak's multiple comparisons). N ≥ 2,016 genes.
[C-D] Pair-wise comparison of transcriptional burst frequencies in the indicated KO populations in unperturbed conditions (D) or 2 days after exposure to stress (C). Each dot is a gene.

In line with the above data, epigenetically disrupted populations also exhibited increased resistance to the upregulation of the burst size in stress related genes. Contrary to the substantial upregulation that was observed in control populations, the KO cells exhibited minimal changes during the first 2 days under starvation with only a significant but mild increase at the latest time point [Figures 60 and 62]. This differential behavior in the regulation of stress genes is in line with the detected differences in the expression of stress genes that were detected between KO and control populations during the immediate response to stress [Figure 53A].

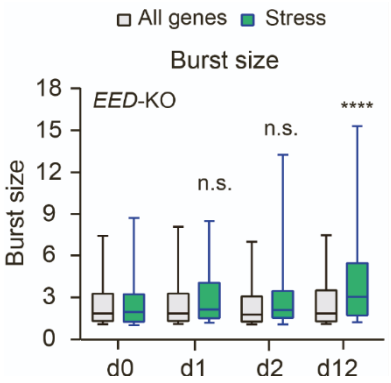


Figure 62: Resistance of *EED*-KO cells to the increase in burst size in stress genes

Distributions of transcriptional burst size for all or stress genes in *EED*-KO cells at the indicated time points under starvation. P-value from two-way ANOVA, followed by Sidak's multiple comparisons (****p < 0.0001). N ≥ 2814 (all) or 57 (stress) genes.

4.3.5.4 Selection of secondary adaptive gene signatures under chronic nutrient starvation

Failing to mount a strong stress response that subsequently leads to the induction of apoptosis increases the probability of survival in the presence of short transient challenges. Additionally in cases of chronic stress, as in the case of sustained nutrient starvation under examination here, it provides valuable time (increased number of surviving cells x increased time under stress) that enhances the probability of acquiring secondary adaptive traits to combat the stressor. In line with that, inspection of epigenetically disrupted cells after 12 days under starvation revealed the presence of a fraction of cells that express glutamine synthetase (*GLUL*) which is responsible for the *de novo* production of this non-essential amino acid (Bott et al., 2015). Other genes known to compensate for glutamine shortage such as *ASNS* (Zhang et al., 2014) and *SLC1A3* (Tajan et al., 2018) were also upregulated [Figure 63A]. Notably, their expression pattern was variable among cells, ranging from cells expressing all three genes to others that express only one. It is worth mentioning that the absence of single genes from individual profiled cells in single-cell transcriptomics experiments can be the result of the stochastic Poison sampling from the examined population. However, no correlation was observed between the expression of the adaptation genes and the overall capture efficiency, as assessed by the total number of detected UMIs or genes in these cells [Figure 63C]. Thus, the mixed expression pattern can be an indication that individual cells employ various strategies for long-term survival [Figure 63B]. The same genes were also detected in the few surviving control cells, but they were significantly less abundant, in line with the overall lower fitness of the population [Figure 63A].

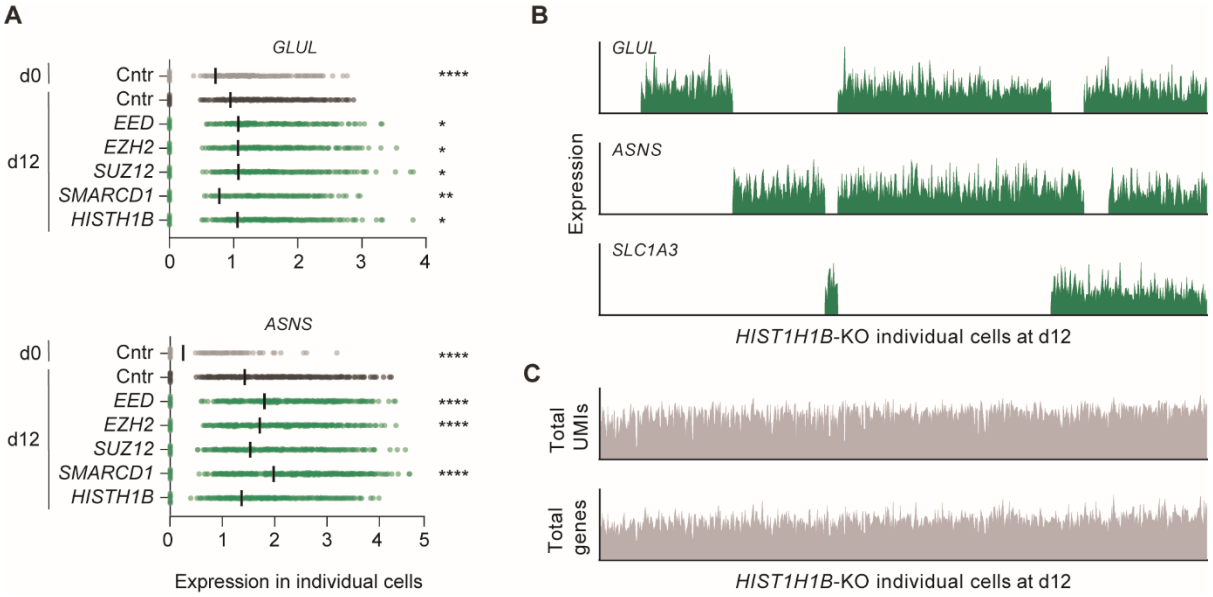


Figure 63: Expression of adaptation genes in epigenetically disrupted cancer cells under chronic nutrient starvation

[A] Expression levels of the indicated genes in individual cells in the indicated samples. Lines indicate the median values. *p < 0.05, **p < 0.01, and ****p < 0.0001 relative to d12 control cells (Kolmogorov-Smirnov test) 492 ≤ N ≤ 664 cells.
[B] Expression levels of the indicated genes in *HIST1H1B*-KO cells at d12 under nutrient starvation. Each line represents the expression level in a cell. N = 664
[C] Total number of captured unique molecular identifiers (UMIs) or genes in *HIST1H1B*-KO cells at d12 under nutrient starvation. Each line represents an individual cell. N = 664

Altogether, the above data demonstrate that epigenetically disrupted cells resist the global alterations of transcription burst properties that take place in response to nutrient starvation, providing a mechanistic explanation for the inert phenotype that they exhibit. This resistance lowers the probability of cell death and concomitantly increases the chance of acquiring secondary traits that can combat the challenge.

4.4 Conclusion

In chapter 4 of this thesis, I have presented step by step the approach that I followed to interrogate the potential cellular mechanism that mediates the stress-dependent selective advantage of epigenetically disrupted cells. In brief, the models explored are a) genetic diversity, b) transcriptional diversity c) phenotypic plasticity and d) phenotypic inertia [Figure 31].

To dissect among them, I employed a combination of experimental approaches ranging from reversible administration of chemical compounds that inhibit epigenetic function to live-cell imaging of metabolic states under stress and finally single-cell transcriptomics in control and epigenetically impaired populations growing under nutrient deprivation [Figure 64].

The key findings presented in this chapter are summarised below [Figure 64]:

- ✓ The survival advantage observed upon epigenetic deregulation is not genetically encoded, as demonstrated by the loss of the stress-resistant phenotype upon withdrawal of drugs targeting specific epigenetic functions [Figure 32].
- ✓ Nutrient deprivation induces strong transcriptomic alterations involving upregulation of stress signatures and downregulation of highly expressed fitness genes [Figure 37].
- ✓ Under chronic starvation, significant phenotypic diversification is observed with the emergence of a mixture of stressed and fit subpopulations [Figures 49, 51 and 52].
- ✓ Epigenetically disrupted cells do not exhibit increased cell state transitions in response to stress as indicated by time-lapse imaging experiments following the metabolic state of individual melanoma cells [Figure 43].
- ✓ Dissection of subpopulation heterogeneity failed to detect subsets of epigenetically disrupted cells that express genes capable of providing adaptation to starvation [Figures 51 and 52].
- ✓ Disruption of the epigenetic control leads to minimal alterations in transcriptional variance [Figure 55].
- ✓ The data obtained from live-cell imaging and that single cell transcriptomics do not support a system where phenotypic plasticity and/ or bet-hedging mediate the stress-resistant phenotype.
- ✓ Epigenetically disrupted cells demonstrate increased tolerance to stress, by adopting a milder phenotype in response to nutrient starvation [Figures 39, 42, 43 and 53].
- ✓ Modelling of the transcriptional burst kinetics reveals a substantial decrease in the burst frequency of fitness genes and a concomitant increase in the burst size of stress genes providing a mechanistic explanation for their downregulation and upregulation respectively [Figures 58, 59 and 60].

- ✓ Epigenetically deregulated cells resist the above stress-induced changes in transcriptional burst properties, thus failing to mount a maximal stress response that would lead eventually to cell death [Figures 61 and 62].

Collectively, the data discussed in Chapter 4 support the notion that phenotypic inertia is likely the cellular trait that underpins the fitness advantage of epigenetically deregulated cells under unfavorable conditions. In molecular terms, inertia is conferred by the inability of KO cells to mount an efficient stress response that ultimately leads to apoptosis and cellular death.

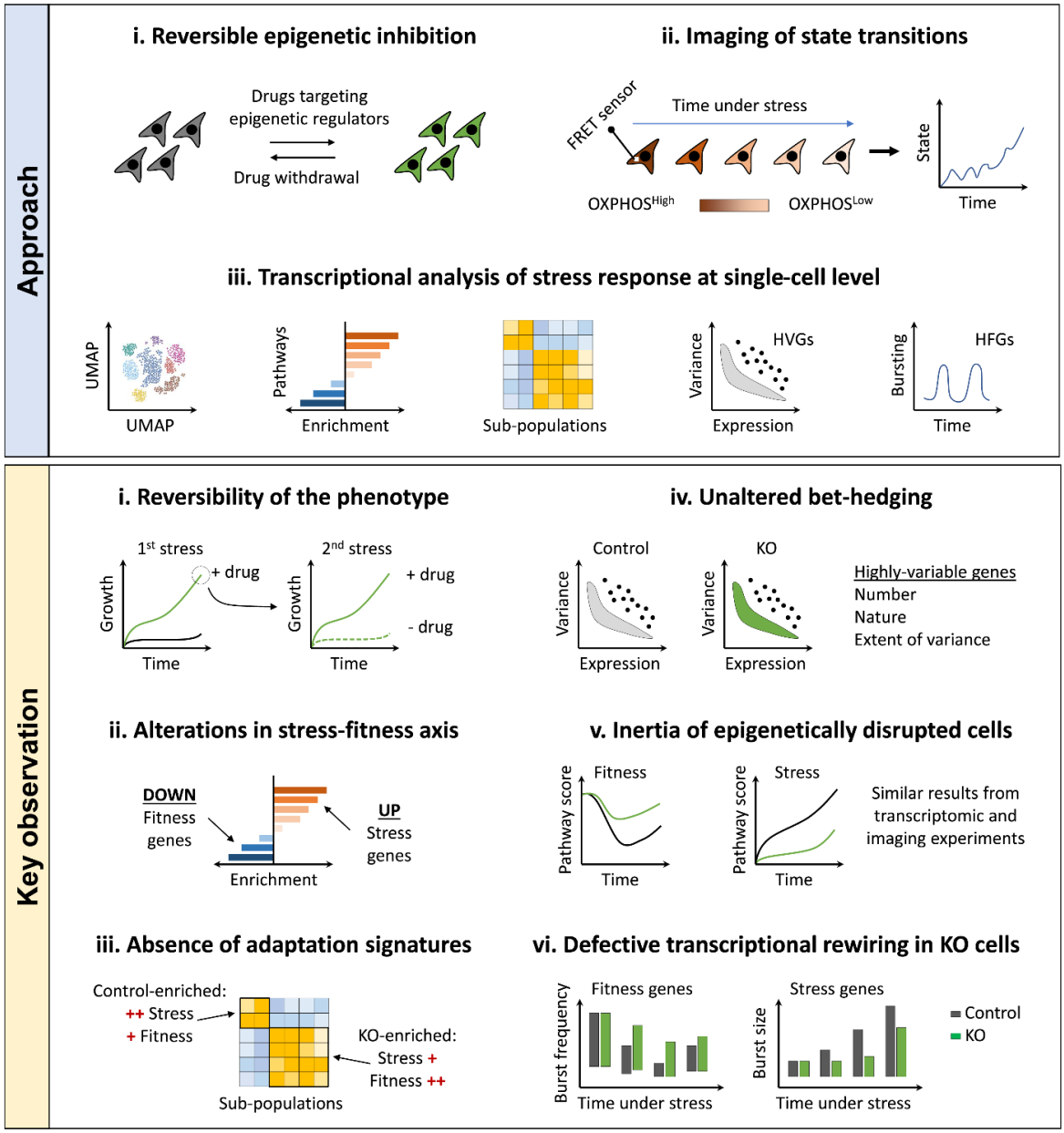


Figure 64: Summary of the experimental methodologies and key observations reported in Chapter 4

Chapter 5. Discussion

5.1 Overview

Cancer is an evolutionary disease that is shaped primarily by natural selection acting upon genetic and epigenetic diversity, to promote the gradual acquisition of favorable phenotypic traits that collectively drive disease progression (Vendramin, Litchfield & Swanton, 2021; Hanahan, 2022). Epigenetic regulators are one of the gene classes that get preferentially disrupted during the later stages of tumorigenesis, yet the functional impact of this disruption remains purely understood (Jamal-Hanjani et al., 2017; Dentro et al., 2021). What is considered a "favorable" selectable trait can vary during tumorigenesis and is influenced by the tumor microenvironment, which is highly heterogenous and dynamic (Lipinski et al., 2016). Within this hostile setting, cancer cells must find ways to withstand and survive stress and subsequently drive tumor progression by acquiring adaptive behaviors. Based on the established role of epigenetic regulators as mediators of response to external stimuli, in this PhD thesis, I have explored the potential relevance of environmental stress in the selective advantage upon their inactivation.

The experimental approach presented in this work consists of a combination of various methodologies, ranging from large-scale arrayed fitness assays, *in vivo* mice models, live-cell imaging, transcriptomic analysis at the single cell level and modelling of various properties of transcriptional regulation and output. In Chapter 3 I demonstrated that disruption of multiple epigenetic regulators in diverse cancer cell lines results in broad survival advantage in distinct metabolic stress conditions, relevant to cancer cells [Figure 30]. In Chapter 4 I investigated various cellular traits that could underlie the stress-specific resistance and propose that the inability of cancer cells to respond to their stressful conditions (Inertia) is the favorable phenotypic trait that is selected [Figure 65]. For a more thorough presentation of the experimental observations made in this thesis please refer to the respective summaries at the end of each results chapter (Sections 3.4 and 4.4).

In this chapter, I present the key points of my study, synthesize observations from both chapters and discuss their implications on various aspects of cancer evolution. On top of that, I explore their limitations and suggest future experimental avenues to further characterize, strengthen and broaden the conceptual ideas proposed in my PhD thesis.

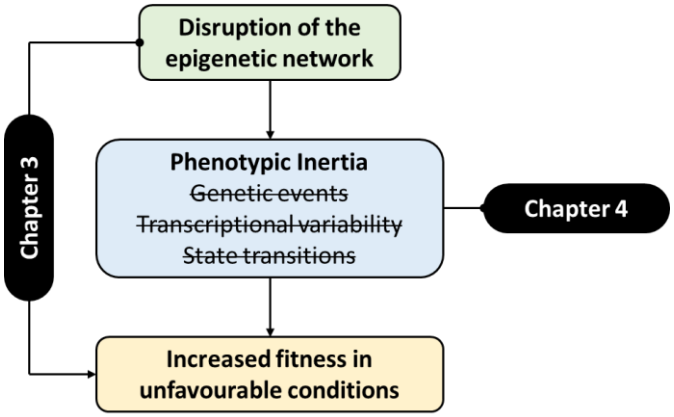


Figure 65: Schematic overview of the conceptual advances made in this PhD thesis

5.2 Broad stress-specific survival advantage upon epigenetic deregulation

5.2.1 Exploring the phenotypic relationships from the large-scale fitness assays

5.2.1.1 Phenotypic commonalities

Inspection of the phenotypic behaviors from the large-scale fitness assays in melanoma and lung cancer cells revealed that i) multiple KO populations demonstrated increased survival under stress ii) the widespread fitness advantage is shared across different stress conditions and cancer types and iii) the inactivated genes that confer advantage belong to all examined functional classes of epigenetic regulation [Figures 20 and 25]. These qualitative and quantitative aspects of the phenotype (number and functional classification of the fit KO populations) argue against the importance of specific epigenetic families and catalytic activities on the phenotype and rather propose that epigenetic deregulation at the network-level converges to a common favorable trait (which in the light of the investigations presented in Chapter 4 is phenotypic inertia). This notion is in line with observations from mutational studies within cancer patients that have revealed non-preferential targeting of multiple genes across the epigenetic families (Brennan et al., 2013; Shen & Laird, 2013). To keep the focus on potential similarities among the detected phenotypes throughout my PhD work, I have explored phenotypic and transcriptomic behaviors in subsets of epigenetically disrupted cancer populations that a) demonstrated varying extents of fitness advantage upon stress and b) are KO for diverse regulators within distinct catalytic activities [Figures 21, 22, 23, 26,33, 36 and 45].

5.2.1.2 Quantitative and qualitative *differences* within the dataset are expected

Aside from the aforementioned commonalities, that served as the conceptual basis of this study, in the fitness assays there were also differences detected across all dimensions of the dataset (e.g. KO populations with varying phenotypes between stress conditions or cancer models). Understanding this variability is important as in most cases it is expected, thus not weakening the generality of the survival advantage as described above.

KO populations that demonstrated fitness neutrality

One potential explanation is that these genes are not expressed in the cancer cell lines, thus no effect is expected upon their perturbation. Indeed, the KO populations within the dataset that did not exhibit any fitness deviation in all the conditions examined, were characterized by significantly lower mean expression compared to the ones that deviated from the norm at least in one stress condition [Figure 18]. Another explanation is that these genes are expressed within the lines, but the system can buffer their perturbation with no detectable phenotypic deviation. Although such robustness can be conferred via multiple mechanisms (Masel & Siegal, 2009), the most straightforward concept is the presence of gene duplications (paralogues), that share sequence similarity and can be characterized by functional redundancy. The methyltransferase paralogues *SUV39H1* and *SUV39H2* that are responsible for the deposition of methyl marks on H3K9 within heterochromatic regions and are characterized by a degree of redundancy demonstrated fitness neutrality across the fitness assays (Table 8) (O'Carroll et al., 2000). Similar behaviors were also observed in other pairs of paralogues like *ARID3A* and *ARID3B*. This is just an indication and integration of this dataset with future explorations regarding the robustness within the epigenetic regulatory network will shed light on the interplay between perturbations, robustness and the observed phenotypic deviation.

KO populations with different phenotypes across stress conditions.

Upon epigenetic deregulation, there was a significant directionality observed in the fitness, mainly towards survival advantage [Figures 20 and 25]. In rare cases, there were classes of genes that exhibited survival disadvantage in a single stress condition. Notable examples in this category are the *CBX1/3/5* and *BRD2/4/8* that both demonstrated a hypersensitive phenotype upon nutrient deprivation. These data suggest a direct link between the respective epigenetic functions and the response of cells to limited nutrients. In large-scale fitness

assays, there is an expected trade-off between accuracy in the detection of a single phenotype and scale. Although reproducibility was high in follow-up validation experiments [Figures 21 and 25], it is still possible that single KO populations displayed falsely no phenotype. It is important to remind that the cancer cell models used in this study display an editing efficiency of approximately 70% [Figure 10], thus providing another technical reason why milder phenotypes may have fluctuated around fitness neutrality, especially in the less sensitive acidic conditions.

Phenotypic differences across cancer models.

Although a significant correlation was detected in the expression profile of epigenetic regulators among melanoma and lung cancer cells, there was also cancer type specific expression of various genes. Thus, the differential set-up within the epigenetic network in each cancer cell model predicts the presence of potential cancer-type specific hits. For example, *HIST1H1B*, *HIST1H1D* and *NAP1L5* were found to be expressed solely in melanoma cells. Indeed, their respective KOs conferred a survival advantage in melanoma cells and not in the lung cancer model. As stated in section 3.3.2.2, to avoid the presence of inflated ratios during the calculation of the stress-specific fitness, KO populations with substantial disadvantage in the unperturbed conditions were discarded from the analysis. KO populations for *CHAF1B* *DDX11*, *HDAC3*, *HMGA1* and *SMARCB1*, demonstrated fitness advantage in melanoma cells, but were discarded in the lung cancer fitness assay. Finally, technical noise stemming from the high-throughput nature of the large-scale assays and the editing efficiency may also account for differences in this dimension of the dataset.

So far, I have focused on single KOs in epigenetic regulators and have discussed reasons for the detection of varying phenotypes within the dataset. Between the melanoma and the lung cancer cells, there was also a significant difference detected in the overall numbers of KO populations exhibiting survival advantage under stress (91 and 38 in MEXF 2090 and LXFL 1674 respectively). The most plausible explanation for this variability is the slow proliferation rate that characterizes the lung cancer cell line, which results in a smaller dynamic range of the detected phenotypes. On top of that microscopic inspection of the NSCL model, revealed significant levels of cell death in the baseline conditions (data not shown). Finally, upon epigenetic deregulation, a significant shift towards slower proliferation was observed in many KO populations in unperturbed conditions. Thus, it is possible that the inherent properties of this model, render them more sensitive to epigenetic deregulation per se, masking potential beneficial effects that this disruption could have across other modalities (like stress

resistance). The interplay of epigenetic deregulation with other existing cellular properties should be characterized more systematically (see subchapter below).

Overall, technical factors inherent to large-scale assays, specific roles of single genes in stress conditions and the structural and functional properties of the epigenetic network across cancer types predict the presence of variability within the dataset.

5.2.1.3 Limitations and future experimental directions

One potential limitation of my thesis is that most of the experiments were performed under nutrient starvation or environmental acidification, both of which constitute metabolic stress conditions. However, these stressors have been shown to induce antithetic effects on cancer cell metabolism, thus the observed stress resistance is expected to emerge from a general phenomenon rather than a specific rewiring in the metabolic apparatus of the cells (Yoo et al., 2020). In line with that, the transcriptomic analysis of the response to nutrient starvation revealed alterations in largely generic pathways like proliferation, inflammation, hypoxia and apoptosis [Figure 37]. Analysis of expression levels in representative genes revealed similar alterations in melanoma cells growing under acidic conditions [Figure 38]. Finally, global alterations in fitness and stress signatures have been previously reported in response to various hostile conditions, ranging from hypoxia to targeted therapy (Emran et al., 2018; Gameiro & Struhl, 2018). The identification of defective transcriptional rewiring in the stress-fitness axis as the common molecular underpin of the stress-resistant phenotype, predicts that the widespread survival advantage upon epigenetic deregulation can be also prevalent in other conditions relevant for cancer cells.

A second limitation of the study stems from the cancer cell models used. Both are clonal lines that are characterized by specific driver mutations in RAS and p53. Thus, extrapolating the observed broad survival advantage to other cancer types needs caution. Follow-up experiments by Fabrizio Simeoni (Scaffidi lab) demonstrated that diverse epigenetic inhibition in PDX models from the colon, pancreas and bladder enhanced the survival of cancer cells under nutrient starvation, thus providing an additional line of evidence that the phenotype is not dependent on the cell of origin (Note that these PDX models are driven by the same oncogenic mutations) (Loukas et al., 2023). Variability in other systematic factors like oncogenic drivers, signalling pathways, differentiation status and genomic instability is likely to significantly affect the detected phenotype upon disruption of the epigenetic machinery (Shen & Laird, 2013). To start dissecting such dependencies, similar approaches employing pharmacological inhibition of diverse epigenetic regulators within cancer cell lines characterised by distinct underlying properties could be performed. Next, to thoroughly dissect

the contributing factors, large-scale fitness assays in cancer models that vary in all the systematic factors should be considered. Although the *arrayed* format of the fitness assays proved powerful in detecting even subtle phenotypes [Figures 17 and 21] and revealing patterns within the dataset [Figures 20 and 25], the execution is quite laborious. An alternative approach here would be to perform *pooled* CRISPR screens in various PDX models under stress. The technical challenge resides in a) the reduced sensitivity towards subtle phenotypes and the fact that different lines are expected to have varying dependencies and modes of response to stress conditions thus potentially limiting the ability to directly cross-compare qualitative and quantitative phenotypic aspects among them. Thus, results should be handled under the prism of network-level effects with the focus being on identifying advantageous phenotypes that are conferred by mutations in many genes (quantitative aspect) that belong to various epigenetic families (qualitative aspect).

5.2.2 Interpretation of cellular phenotypes driven by phenotypic inertia

Aside from the binary investigation of the stress response (fitness neutrality vs fitness advantage), various experiments within Chapter 3 explored additional parameters of the stress response like the fraction of surviving cells, the trajectory over time and how epigenetic deregulation may affect them. In brief, those experiments revealed that the survival advantage of epigenetically disrupted cancer cells correlated with the ability of cells to sustain an increased proliferation-to-apoptosis ratio [Figure 22] and an increased number of stress-resistant colonies [Figure 23]. However, a close comparison of the phenotypes among the different KO populations, revealed cases that diverge from the above general notions. For example, *HIRA*-KO cells presented more stress-resistant colonies, but ultimately lower cell count compared to *SMARCD1*-KO populations [Figures 21A and 23B]. Additionally, following the trajectory of dozens of KO populations under stress demonstrated multiple cases where the relative ranking of the detected fitness was altered over time [Figure 21B]. Thus, it is evident that the extent of the survival advantage in the KO cells is shaped by multiple factors like the capacity of cells to combat cell death in the early response to stress along with additional secondary mechanisms that emerge and dictate long-term growth under stress.

Notably, the identification of inertia as the mechanism behind the stress-specific phenotype aligns with this view of the system and allows for the presence of the aforementioned “discrepancies”. Upon stress application, the inert phenotype decreases the potential of immediate cell death thus increasing at the population level the chance of acquiring secondary traits that can sustain growth under stress. This acquisition is largely stochastic and can vary in regard to timing (if and when the adaptive event happens) and its nature (how it affects

proliferation under stress). As proof of principle, the transcriptomic analysis detected at the latest time point (d12) significant cell-cell variability in the expression of various genes that have been shown to compensate for glutamine starvation [Figure 63]. Finally, this stochasticity can also explain the observations from the *in vitro* competition assays where variability in the response was observed even among replicates of the same KO population [Figure 26B].

Considering the variability within the explored phenotypes, further experiments are required to fully characterize the phenotypic response to stress of epigenetically disrupted cancer cells. The incorporation of barcoding and lineage tracing methodologies would provide important insights, regarding the clonality of the response and how the system evolves over time. It is worth reminding that even within the KO populations a significant fraction of cells succumb to stress [Figure 22B]. What drives this variability in the response? Are there pre-existing heterogeneous traits within the phenotypically inert cells that can tip the balance between survival and death upon stress? Utilizing barcoding systems like Rewind (Emert et al., 2021) and CaTCH (Umkehrer et al., 2021) that enable the detection of surviving clones and the subsequent interrogation of the traits of the ancestral clone in the unperturbed conditions will shed light towards this direction.

5.2.3 Expanding the fitness relationships in the *in vivo* setting

Most of the investigations presented in my thesis were performed *in vitro*. This allowed for the i) elimination of various potential confounding properties of the system ii) direct interrogation of the relationship between epigenetic deregulation and survival under stress iii) execution of complex experiments that would have been challenging *in vivo*. Having established that inactivating mutations in diverse epigenetic regulators can increase the survival of cancer cells under stress, the next step was to explore if this link holds true in the much more complex setting of evolving tumors.

Initial evidence in support of this notion, comes from the *in vivo* competition experiment, where *EZH2*-KO cells outcompeted control populations in all tumors where the microenvironment was manipulated to promote further nutrient scarcity [Figure 28]. One potential limitation of this experimental setup arises from the fact that within the mixture of cells that are injected into mice, a fraction of cells is already knock-out for the epigenetic regulator. This creates the possibility that part of the detected phenotypic deviation could be independent of the ability of cells to respond to environmental stress and could be attributed to other traits like increased capacity to initiate tumorigenesis or increased proliferation rate *in vivo*. However, the dichotomy among the untreated tumors (*EZH2*-KO enrichment in two vs fitness neutrality in the rest) dismissed this scenario [Figure 29]. On top of that, the fact that *EZH2*-KO enrichment

correlated with lower nutrient availability (as indirectly inferred by pS6 staining) confirms that the detected phenotype is dependent on the relationship of the cells with the TME [Figure 29]. Finally, it demonstrates that even within replicate models of tumorigenesis, the baseline levels of nutrient availability can substantially affect the selective forces and thus the detected phenotype.

The above observations are encouraging but limited only to one KO population. Further *in vivo* competition assays are needed to strengthen the detected *in vivo* phenotype. A better future approach would be to perform such an experiment with an inducible setup, where the KO is initiated after the formation of a primary tumor, thus more faithfully mimicking the scenario of subclonal expansion [Figure 66A]. Additionally, using KO and control melanoma cells (MEXF 2090) stably expressing the GFP and mCherry fluorophores [Figure 26] would allow for their direct detection within the tumor core and assessment of their spatial distribution relative to other components of the TME. This approach combined also with spatial transcriptomics (Lewis et al., 2021), could provide powerful insights into the biology behind the spatial distribution and expansion of epigenetically deregulated cancer cells in evolving tumors.

Going a step further, it would be interesting to explore signs of the *stress-epigenetics* relationship in cancer patients. This can be done by a combination of computational and experimental methodologies, with the core principle being to detect relative enrichment of mutations in epigenetic regulators within differentially stressed tumor samples (similar concept to the *in vivo* competition assays in mice). It is well established that cancer types are characterized by overall different levels of hypoxia, which subsequently generates nutrient scarcity and environmental acidification (Bhandari et al., 2019). Thus, one potential avenue is to look for a positive correlation between the number of subclonal mutations in epigenetic regulators and the overall hypoxic score of those tumors [Figure 66B]. An additional computational approach, that takes into account various interpatient confounding factors, is to focus on tumors with available multi-region sequencing data (Jamal-Hanjani et al., 2017), stratify the subclones based on the expression of various stress signatures (Barkley et al., 2022; Baron et al., 2020) and look for enrichment of epigenetically deregulated cells [Figure 66C]. A reciprocal approach can also be applied here. Although this approach is more straightforward, it requires a significant number of specimens that have been profiled both at the genetic and transcriptional levels. Finally, an experimental equivalent to the above methodologies is to select patients with mutations in epigenetic regulators and co-stain with markers of stress, to directly examine the spatial proximity between these deregulated cells and stress signatures in the *in vivo* setting [Figure 66D].

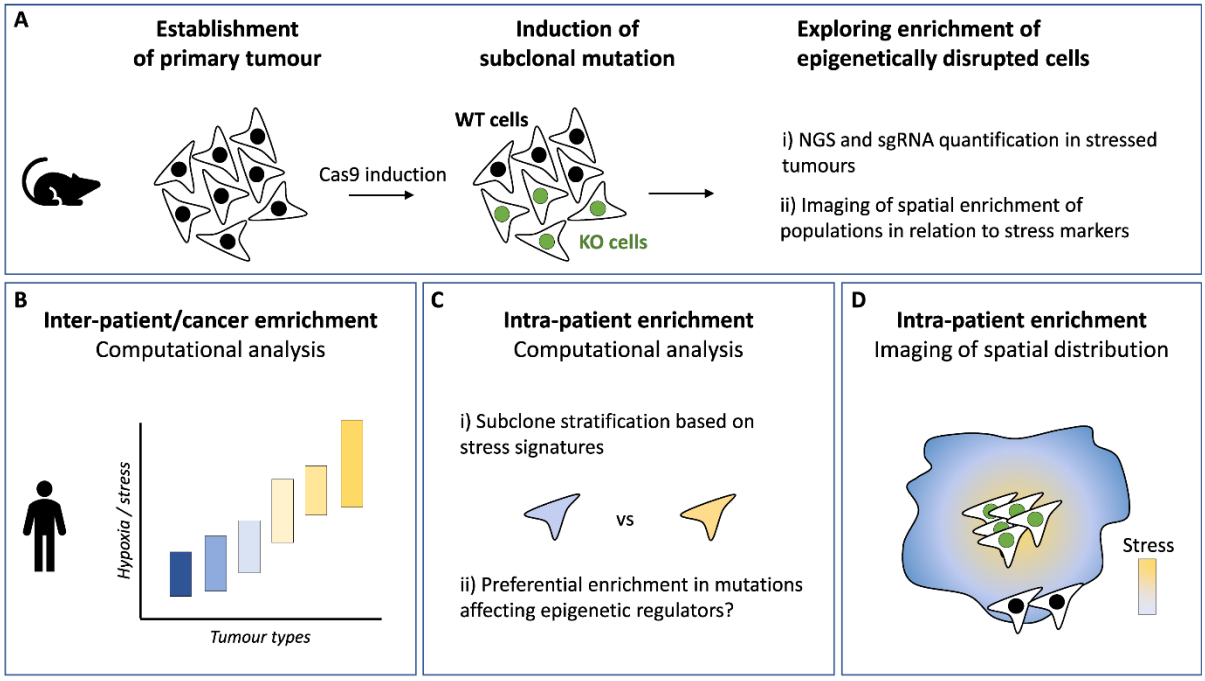


Figure 66: Proposed experimental directions to strengthen the notion that epigenetic deregulation alter the sub-clonal expansion of cancer cells in evolving tumors

5.3 Transcriptional response to stress

The profiling of mean expression, variance and bursting properties of genes along the time trajectory, allowed for a multifaceted investigation of how melanoma cells respond to stress [Figure 67]. Nutrient starvation induces a substantial number of transcriptional alterations that gradually accumulate over time [Figure 48A]. In the early response to stress (d2 vs d0) highly expressed genes involved in proliferation and cell cycle control get downregulated whereas pathways indicative of stress (e.g. Hypoxia, NFKB, Apoptosis) get upregulated [Figures 37, 48B and 49]. In the long-term response (d12) control cells also demonstrate a strong inflammatory phenotype, in line with previous reports further confirming the validity of the transcriptomic analysis (Gameiro & Struhl, 2018) [Figure 49]. Interestingly, I observed at the latest time point a partial recovery in fitness related signatures suggestive of adaptation at least in a fraction of cells [Figures 48C and 49]. The above evolutionary conserved response has a dual role as it allows for the conservation of cellular energy and biomass along with the upregulation of stress related genes to promote short-term survival under stress. It is currently unclear if the detected alterations in fitness and stress related genes are interdependent. Finally, another observation within the dataset was the transcriptomic diversification in response to stress, which resulted at the latest time point in control cells that exhibited significant variability along the stress-fitness axis [Figures 47, 49 and 51]. It is worth

mentioning that the above analysis of the transcriptomic response was focused on the context of glutamine starvation. However, various evidence supports its generality. For example, similar alterations in fitness and stress related genes were observed in acidic conditions [Figure 38]. On top of that similar alterations have been previously reported in other contexts including diverse metabolic stressors and therapy (Emran et al., 2018; Gameiro & Struhl, 2018). Thus, it is plausible that the above transcriptomic alterations can extend to additional hostile conditions, a concept that needs to be experimentally validated.

One interesting observation is that fitness and stress related genes are characterised by distinct bursting properties. Specifically, highly expressed fitness genes exhibit elevated burst frequency, while stress genes demonstrate lower frequencies compared to the norm [Figures 59 and 60]. This dichotomy likely reflects the evolutionary shaping of hard-wired and dynamic properties that dictate the bursting behavior of the respective expressed loci. Such factors within fitness and stress related genes have been previously reported in other contexts and cellular models, with the most notable example being the preferential presence of TATA box elements within the promoter regions of stress related genes that results in increased expression noise (Bar-Even et al., 2006; Newman et al., 2006; Larsson et al., 2019; Ochiai et al., 2020). This relationship was also confirmed by the variance analysis presented in this thesis [Figure 54].

The fact that starvation induces a strong reduction in global burst frequencies [Figure 58], while size remains unaffected, provided a mechanistic framework for the coordinated down-regulation of functionally related fitness gene sets, that spam multiple chromosomes and are under the regulations of distinct GRNs. The low frequency of stress genes makes them less sensitive to the induced reduction in frequency and their upregulation seems to correlate with a gradual increase in burst size, potentially reflecting the stable binding of stress induced transcription factors. Such proteins like ATF3, ATF4 or DDIT3 were significantly upregulated upon stress (data not shown) and their functional relevance regarding the bursting alterations remains to be elucidated. Overall, the above data reveal a dichotomy in traits within stress and fitness related genes that enables their differential regulation. Similar selectivity has been also observed at the translational level, via alterations in the dynamics of tRNAs (Torrent et al., 2018), revealing a multi-layered evolutionary regulation across distinct layers of gene expression in response to stress.

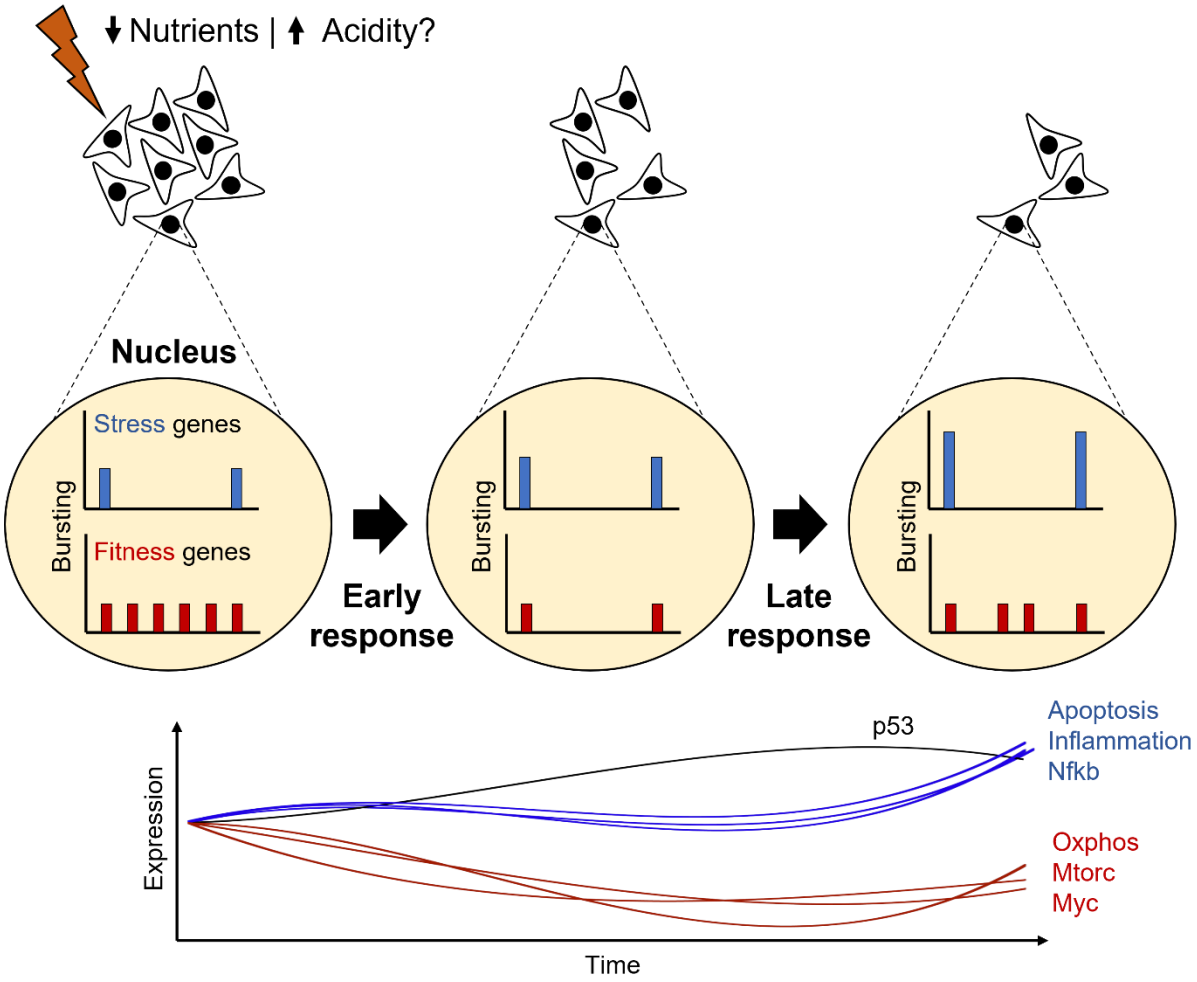


Figure 67: Schematic representation of the transcriptomic response of melanoma MEXF2090 cells under nutrient deprivation.

The lines have been arbitrarily selected to start from the same position within the Y axis to aid the visual inspection of the trends over time.

5.4 The stress resistant phenotype is not mediated by genetic events, transcriptional variance or state transitions

Having established the link between epigenetic deregulation and increased survival under stress, the main focus of Chapter 4 was to explore the presence of underlying traits that could mediate this phenomenon. In brief, I investigated if the stress resistance of the KO populations was conferred by genetic events, transcriptional variance, state transitions or inability to alter their phenotype in response to stress [Figure 31].

Multiple lines of evidence argue against the importance of mutations as drivers of the stress-tolerant phenotype upon epigenetic deregulation. First, the clonal nature of the PDX lines

eliminates the presence of pre-existing heterogeneity at the genetic level. On top of that, whole exome sequencing of spontaneously adapting parental lines to starvation failed to identify clonal or shared subclonal driver mutations [Figure 13]. The complete loss of the favorable phenotype in all populations, upon withdrawal of the epigenetic inhibitors, is the strongest hint against the genetic nature of the favorable trait [Figure 32]. It is theoretically possible that the established mutation confers a phenotypic advantage that is dependent on the existing deregulated epigenetic state. This unlikely scenario is not supported by the growth kinetics of KO populations under stress. A model where a rare genetic event (pre-existing or *de novo* acquired) is favorable and selected over time would require weeks to drive detectable alterations at the phenotype (cell count). Contrary to that prediction, pharmacological or genetic inactivation of epigenetic regulators led to significant fitness leaps after only a few days under starvation, as determined in various fitness assays [Figures 14, 20, 21 and 45].

The second scenario that was assessed is if epigenetic deregulation can affect the transcriptional variance of cells, priming them for a better response to stress, a concept analogous to bet-hedging in unicellular organisms. In this study, transcriptional variance was assessed at the gene level. Inference of variability was robust as the defined cohort of highly-variable genes was enriched for stress and plasticity related signatures [Figure 54]. This is in line with the existing notion, that across species selective forces have shaped the regulatory properties of these cells in order to minimize variance within housekeeping genes and maximize heterogeneity within stress related signatures (Bar-Even et al., 2006; Newman et al., 2006). No difference in the number, extent or nature of the HVGs was detected among the KO and the control populations. My analysis was restricted to the unperturbed conditions (day 0), with the idea being to identify pre-existing variability that could prime cells for a better stress response. Towards this direction, inspection of the transcriptional variance within *GLUL, ASNS and SLC1A3,* genes that have been reported to compensate for glutamine starvation (Zhang et al., 2014; Tajan et al., 2018), revealed no differences among control and epigenetically deregulated cells (data not shown). Other forms of variability within the population were also assessed (Grün, 2020), but the results were inconsistent and sensitive to subtle parameter changes thus limiting the power to draw conclusions from them (data not shown). In line with the observation in my PhD thesis, direct assessment in an experimental model where heterogeneity has been shown to be relevant for survival under stress, did not detect the emergence of favorable phenotypes upon disruption of the epigenetic network (Torre et al., 2021).

An alternate concept to transcriptional variance is cellular plasticity, a term heavily misused in the cancer literature (Mills, Stanger & Sander, 2019). Here plasticity has been examined under the prism of state transitions, by utilizing a FRET based biosensor that can indirectly monitor

the OXPHOS state of cancer cells, which was demonstrated to be responsive to stress. Live-cell imaging of individual EZH2i and DMSO melanoma cells did not reveal any differences in the state variability during the initial response to starvation [Figure 43]. Although this analysis should be extended ideally to other KO populations and additional monitored transcriptional states, it provides initial evidence against the increased plasticity of cancer cells upon epigenetic deregulation. If indeed the KO cells were more robustly transitioning towards a favorable (transcriptional) state, then it is expected that this state would be selected over time. Dissecting the sub-population heterogeneity at the latest time point (day 12) did not detect any subsets of epigenetically disrupted cancer cells with adaptive signatures but rather cells with milder alterations in the stress-fitness axis [Figure 52]. These differences were not a result of recovery after reaching a full stress maxima, as following the fitness and stress signatures along the time trajectory revealed increased resistance of epigenetically disrupted cancer cells in response to stress [Figure 53]. Those observations are in accordance with the behaviors observed in EZH2i cells in the experiments monitoring the metabolic states [Figures 42 and 43].

Overall, observations from fitness assays, live cell imaging and the subsequent transcriptomic analysis provided multiple evidence in favor of phenotypic inertia as the mediator of the advantageous phenotype in KO cells [Figures 39, 42, 43, 53, 61 and 62]. It is essential to state that the data presented in this thesis do not imply that epigenetic deregulation does not lead to genomic instability or plasticity or transcriptional variance. On the contrary, they propose that there are specific contexts and time scales during cancer evolution (like the examined environmental challenges) that disruption of the epigenetic network confers a survival advantage through alternate non-genetic mechanisms like phenotypic inertia.

5.5 Considerations regarding the causes and consequences of phenotypic inertia during cancer evolution

5.5.1 Assessing the causality of the detected transcription burst alterations

The data presented in my PhD thesis (Chapter 4) are in line with the existing literature and suggest that diverse stressful stimuli promote a significant alteration in the transcriptional profile of cells, consisting primarily of the downregulation of fitness related genes and upregulation of stress signatures. Modelling of bursting properties revealed that this downregulation is mediated by a decrease in burst frequency while the upregulation of stress genes is primarily driven by an increase in burst size. It is important to state that follow-up investigations by Fabrizio Simeoni confirmed the predictions from the single-cell

transcriptomics and bursting analysis, which was used as a proxy of global transcriptional activity. In detail, fitness genes exhibited a 60- 90% reduction in nascent RNA levels after two days under starvation accompanied by a significant decrease in bound RNA polymerase II phosphorylated at serine 2 within its CTD domain (this modification marks active RNA polymerases during transcriptional elongation) (Loukas et al., 2023). The proposed model is that the increased resistance of the epigenetically deregulated cells towards those alterations (herein termed transcriptional "numbness") leads to their inert phenotype and thus to increased survival under stress. One major question that arises is if transcriptional numbness is causal to the survival advantage. An alternative scenario is that the KO populations are characterized by a pre-existing cellular trait that primes them for a better stress response, thus any detected alterations in the bursting properties are a passive trait within the system.

Multiple hints along the presented data in this study argue against this scenario. For example, the fact that the stress-resistant phenotype is conferred by mutations in genes across all known epigenetic families argues against the selection of a single "favorable" molecular function (e.g. writing of specific methylation marks is important for the phenotype) [Figures 20 and 25]. Accordingly, the shared widespread advantage in response to two antithetic stressors eliminates the relevance of potential stress-specific favorable traits (e.g. decreased glutamine dependence upon epigenetic deregulation) [Figures 20 and 25]. In the unperturbed conditions, *SMARCD1*-KO and *HIST1H1B*-KO cells are characterized by minimal pre-existing transcriptomic alterations [Figure 34A], thus decoupling the extent of the survival advantage with the presence of any pre-existing transcriptional rewiring. In line with that, the levels of the stress-responsive genes were identical among the examined KO populations before the application of the stress [Figure 39C]. Finally, the most important hint regarding the causality of the phenotype is the nature of the epigenetic regulation itself, as its primary function is to shape the transcriptional output in response to diverse internal and external stimuli.

Nevertheless, it is essential to directly explore the causality of transcriptional numbness. Ideally, this would consist of a perturbation within the system that can affect the capacity of cancer cells to alter their bursting properties in response to stress. Bursting frequency and size have been proposed to be regulated by a variety of fixed but also dynamic traits within the cells, with RNA polymerase being an important factor in the latter category (Rodriguez & Larson, 2020; Tunnacliffe & Chubb, 2020). In other experimental systems pharmacological inhibition of RNA polymerase elongation has resulted in deterministic effects on bursting frequency and size (Ochiai et al., 2020). Thus, one potential experimental set-up would comprise of treating MEXF 2090 cells with inhibitors affecting various steps of RNA polymerase function (initiation, elongation, etc) and assessing the survival under multiple stress conditions like nutrient starvation and environmental acidification. In an alternative

approach, Fabrizio Simeoni genetically inactivated in MEXF 2090 cells *NELFA*, a component of the NELF complex that negatively regulates the elongation of transcription by RNA polymerase II (Aprile-Garcia et al., 2019; Rawat et al., 2021). Strikingly, these KO cells exhibited increased survival advantage under nutrient starvation, in analogy to the behavior in epigenetically deregulated cells, providing important evidence in favor of the causal relationship between transcriptional numbness and the ability of cells to survive under stress.

5.5.2 Mechanistic convergence upon epigenetic deregulation?

Why are epigenetically deregulated cells defective in the transcriptomic response to stress? This is an especially interesting question, considering that the examined regulators are characterised by different functions and catalytic activities. Considering their nature, future experiments should focus on how properties within the chromatin landscape change upon stress. For example, one scenario is that the KO cells fail to remodel the chromatin (histone marks and/or accessibility) around fitness and stress genes thus hindering the expected stress-induced transcriptomic alterations. Significant changes in histones marks (Sharma et al., 2010; Emran et al., 2018) and chromatin accessibility (Shaffer et al., 2017) have been reported in various cancer models that acquire resistance to either environmental stress or therapeutic intervention. However, these alterations are detected in long-term survivors under stress, having undergone significant epigenetic reprogramming, and may reflect alterations that are not important for the early stress response.

In line with that, Marta Milan and Fabrizio Simeoni focused on the initial response to stress and assessed how chromatin dynamics and modifications (via ATAC-seq and Chip-seq respectively) compare in unperturbed conditions and after two days in starvation. Despite the apparent global transcriptional downregulation, the control populations under stress exhibited a striking upregulation of their chromatin accessibility. On the contrary, active or repressive histone modifications remained seemingly unaffected, in line with the pattern observed during the transcriptional downregulation within cells upon UV irradiation (Liakos et al., 2020). Epigenetically deregulated cells demonstrated minimal alterations in chromatin accessibility, providing a first indication that distinct epigenetic activities when mutated converge to transcriptional numbness through the inability of cells to remodel their chromatin landscape in response to stress (Loukas et al., 2023).

5.5.3 Is epigenetic deregulation the only road to inertia?

One important clarification to be made is that transcriptional numbness and phenotypic inertia are two separate concepts. The defective rewiring of transcription in response to stress (numbness), as a result of epigenetic deregulation, is the *molecular* trait that leads to an unresponsive cellular state (Inertia) which in turn is the *phenotypic* trait under selection. This decoupling among molecular and phenotypic properties predicts that perturbations in other regulatory mechanisms within the cell may result in an inert state. For example, signalling pathways and proteins involved in translational control, are both implicated in sensing and responding to environmental stimuli and have been shown to be important for cellular response to stress, including therapy (Lee et al., 2021). This scenario would be in line with the preferential inactivation of signalling proteins that has been previously described during subclonal expansion in evolving tumors (Jamal-Hanjani et al., 2017).

During the large-scale fitness assays presented in this thesis, there were two important signalling genes targeted as a control; mTOR, an evolutionary conserved sensor of nutrient availability and cellular energy and ATM, a kinase that responds to genotoxic stress by promoting cell cycle arrest, DNA repair and apoptosis. Notably, these KOs exhibited increased survival advantage under stress, providing preliminary evidence towards this notion (Table 8). In a non-cancerous setting, cells lacking PTEN, a protein that is commonly mutated in cancer and acts as a negative regulator of PI3K signalling, demonstrated a fitness advantage in conditions of limited nutrient availability (Nowak et al., 2013). In a MYC driven model of neuroblastoma, inhibition of ATF4 primed cells for a better response to starvation, suggesting that deregulation of transcriptomic events downstream of signalling cascades can alter the phenotype under stress (Qing et al., 2012). Finally in cancer models of stress resistance in various stressors, including glucose deprivation, strong alterations in the activity of signalling pathways coincided with alterations in the chromatin state, suggesting a potential crosstalk among these regulatory layers (Emran et al., 2018). Further experimental investigation is required to define perturbations outside the epigenetic regulatory network that can result in an inert phenotype. *Arrayed* fitness assays in MEXF 2090 melanoma cells under various stress conditions, systematically targeting genes involved in signalling and transcriptional regulation would provide powerful insight towards this direction. Similar methodologies have been applied for the exploration of the functional impact of such regulatory layers into other aspects important for cancer evolution, like transcriptional heterogeneity and adaptation to targeted therapy (Torre et al., 2021). An experimental shortcut would be to treat the melanoma cells with various available inhibitors against signalling proteins and assess survival under stress.

5.5.4 Relevance of inertia for other stress conditions during disease progression

In this PhD thesis, the defective transcriptional rewiring of the epigenetically disrupted cells was detected in the fitness-stress axis. In principle, phenotypic inertia is predicted to be relevant in response to any internal or external stressor that acts through transcriptional alterations and is strong enough to tip the balance between fitness and stress related signatures. The strength of the stress (dose x time under exposure) is important, as varying severity may induce alternative behaviors within the cells ranging from minor adaptive rewiring to cell death. In line with that, when performing validation experiments in conditions with varying acidity [Figure 25C], the extent of the observed fitness advantage in epigenetically disrupted cells (specifically in KO populations for either PRC2 subunit), positively correlated with the severity of the stress (as assessed by the extent of cell death and the reciprocal upregulation of stress related genes – data not shown). Alterations in similar pathways and/or global downregulation of transcription have also been reported in a variety of cellular models under stress, including cells exposed to other metabolic stresses, hypoxia, UV irradiation and thermal stress (Emran et al., 2018; Gameiro & Struhl, 2018; Tufegdžić Vidaković et al., 2020; Cugusi et al., 2022). Thus, it would be interesting to explore if the inert phenotype is advantageous within these settings too. Selection of the populations used for the validation experiments [Figure 21] and monitoring of their fitness under stress would provide interesting evidence regarding the generality of inertia within environmental challenges. Of note, the melanoma cells are insensitive to hypoxia, thus examination of the relationship in this specific stress is not possible [Figure 24C].

The inability of cancer cells to respond to unfavorable conditions (a concept to be expanded), suggests that inertia can be a relevant trait for other aspects of cancer evolution, like metastasis. Metastatic dissemination is a complex biological process consisting of multiple steps like local invasion, transport through the bloodstream and finally seeding and proliferation within the recipient tissue (Fares et al., 2020). During this process cancer cells face a multitude of hostile conditions directing them towards cell death. Interestingly, in a genomic study across various solid tumors, targeting enzymes involved in chromatin regulation was shown to be enriched among the metastatic driver mutations (Hu et al., 2020). On top of that, multiple accumulating evidence suggest that non-genetic mechanisms can act as drivers of this process (Chen & Yan, 2021; Gui & Bivona, 2022). Experimentally establishing primary tumors with a mixture of fluorescently labelled control and KO populations and exploration of the nature of the cells within the bloodstream (single or CTC clusters isolated by FACS) could provide initial evidence regarding the ability of these cells to promote local invasion and survive within the bloodstream. On top of that, tail vein assays could further

explore the overall metastatic potential of these cells, including their ability to survive within the new setting and promote secondary growth.

Aside from naturally occurring hostile conditions, one of the most significant struggles that cancer cells face is therapeutic intervention. Induction of a strong inflammatory response and downregulation of various fitness signatures has been also reported as a response to both chemotherapy and target therapy in various cancer models (Daigeler et al., 2008; Komurov et al., 2012; Emran et al., 2018; Rambow et al., 2018; Aissa et al., 2021), thus necessitating the investigation of the potential relevance of inertia to therapy response and resistance. MEXF 2090 cells are $BRAF^{WT}$ driven by oncogenic NRAS mutation. NRAS can exert its oncogenic function through a variety of downstream signalling cascades including the RAF/MEK/ERK, MAPK, and PI3K/AKT pathways (Johnpulle, Johnson & Sosman, 2016). Thus, an appropriate experimental approach would be to compare the fitness of control and KO populations upon inhibition of one of these downstream effectors (or a combination of them). Upon treatment of melanoma cells with PI3K inhibitors, Fabrizio Simeoni observed increased fitness in a panel of epigenetically deregulated populations, thus providing a first line of evidence regarding the potential importance of phenotypic inertia in resistance to therapeutic intervention (Loukas et al., 2023).

In their recent study Torre *et al*. screened various epigenetic regulators for the effect on transcriptional variability and stress resistance to therapy. In contrast to the data presented here, upon epigenetic deregulation, they observed overall fitness neutrality, with only a handful of mutations in epigenetic regulators conferring stress resistance to targeted therapy. Although the two models are fundamentally different in many regards, this phenotype is potentially an underestimation as the pooled CRISPR screen was not run towards saturation. Interestingly within the dataset there were KOs in epigenetic regulators (e.g. KMT2D) that had no effect on the number of “primed” cells but demonstrated increased resistant colonies upon therapy, indicating an alternative mechanism of action compared to transcriptional variance. Phenotypic inertia could fit into this model. Accordingly, Marsolier and colleagues detected a distinct subpopulation in the treatment-naïve breast cancer cells that exhibited increased capacity to withstand therapeutic intervention. These cells exhibited a distinct epigenetic state characterized by alterations in the landscape of the repressive H3K27me3 mark, which is deposited on chromatin by EZH2. Inhibiting EZH2 in these cells increased the resistance of cells to therapy, in an analogy to the increased stress resistance of EZH2-KO cells in our model systems (Marsolier et al., 2022).

Targeting of the epigenetic mechanisms has been proposed as an attractive therapeutic approach, via the elimination of the plastic behaviors that have been observed within cancer cells (Marine, Dawson & Dawson, 2020). In light of the data presented here, caution is needed

regarding this therapeutic avenue, as inhibition of the epigenetic machinery may select for the inert phenotypic which is predicted to be relevant for therapy, thus yielding potential antithetic results. Further studies are required to carefully examine the prevalence, sources, and consequences of phenotypic inertia, as a result of epigenetic deregulation.

5.5.5 Inertia and other models of cancer evolution

As discussed above, further studies are needed to characterize key properties of the inert phenotype and assess its significance in various stages of tumor evolution. An important step forward is to start positioning phenotypic inertia within the existing framework of non-genetic mechanisms that drive cancer cell tolerance. The adaptation of cancer cells during disease progression has been largely studied under the lens of resistance to therapeutic intervention. Aside from genetic events, the selection and/or emergence of cancer cell subpopulations with intrinsic tolerance to stress is emerging as a major non-genetic driver of tumorigenesis (Marine, Dawson & Dawson, 2020). Despite the significant variability across model systems, these drug-tolerant cells are characterized by some shared/overlapping properties. The pre-existing or emerging resistant cells are rare, with some studies positioning their clonality within a range of 1:100-1:10000. Such variability likely reflects both differences within the biological mechanisms involved in various cancer models but also technical variability in the methodologies and timeframes used during the experimental investigation. Nevertheless, the tolerant phenotype is associated with the transition towards a slowly proliferating cell state and is also accompanied by alterations in cell identity ranging from reversion to an undifferentiated stem-cell like state to trans-differentiation towards an alternate cell lineage (Shen, Vagner & Robert, 2020). Contrary to the above characteristics, the increased survival advantage of epigenetically deregulated cells was not a rare phenomenon [Figure 14], it was associated with increased proliferative capacity of cells in the immediate response to stress [Figure 22] and was independent of alterations in stemness [Figure 34]. On top of that in persister cells, activation of stress response signalling and the subsequent upregulation of ATF4 transcription factor mediates the survival of melanoma cells upon targeted BRAF inhibition (Yang et al., 2021). In our model, cancer cells with decreased activation of stress signalling are the prevailing component, suggesting that the survival advantage of epigenetically deregulated cells is dependent on different modalities. Phenotypic inertia renders the cells less sensitive to their hostile setting thus decreasing the possibility of immediate cell death.

5.5.6 Assessing acquired vulnerabilities in inert cancer cells

So far, the emergence of phenotypic inertia has been largely discussed under a positive prism, as it is considered a favorable trait that can provide survival advantage under environmental stress and is predicted to do so also in other stress conditions that substantially affect the fitness-stress axis, like targeted therapy (See section 5.5.4). One interesting question is if the emergence and selection of the inert cellular state is accompanied by the acquisition of novel vulnerabilities that then could be exploited for their elimination. In an analogous manner, stress-tolerant cancer cells, a non-genetic mechanism of therapy resistance, have been reported to acquire vulnerabilities in various pathways ranging from DNA damage response to metabolism (De Conti, Dias & Bernards, 2021).

In preliminary data that are not displayed in this thesis, KO cells in either subunit of the PRC2 complex displayed hypersensitivity to replicative stress after growth and expansion in nutrient starvation. Generalization from such observations needs caution as phenotypes detected within single KOs can reflect gene-level and not network-level effects. Nevertheless, they suggest that stress-induced selection can shape novel properties within cells that are worth further investigation. Considering the defective transcriptional rewiring in epigenetically deregulated cells under stress [Figure 61], an intriguing scenario is that epigenetic deregulation followed by stress-dependent selection of the inert cell state renders cells hypersensitive to perturbations affecting transcription per se. To test this hypothesis, Fabrizio Simeoni cultured control and epigenetically deregulated cells under starvation and isolated the stress-resistant populations after prolonged culture. Treatment of parental and stress-resistant populations with BET and CDK9 inhibitors revealed a hypersensitivity of the latter to transcription inhibitors (Loukas et al., 2023). These data provide initial insights into the acquired vulnerability of phenotypically inert cells and suggest a potential therapeutic avenue towards their elimination. Furthermore, robustness in biological systems to genetic perturbations can be affected by stress conditions. Thus, another intriguing possibility to be explored in the future is that the cells selected under stress become more sensitive to further disruption of their epigenetic machinery.

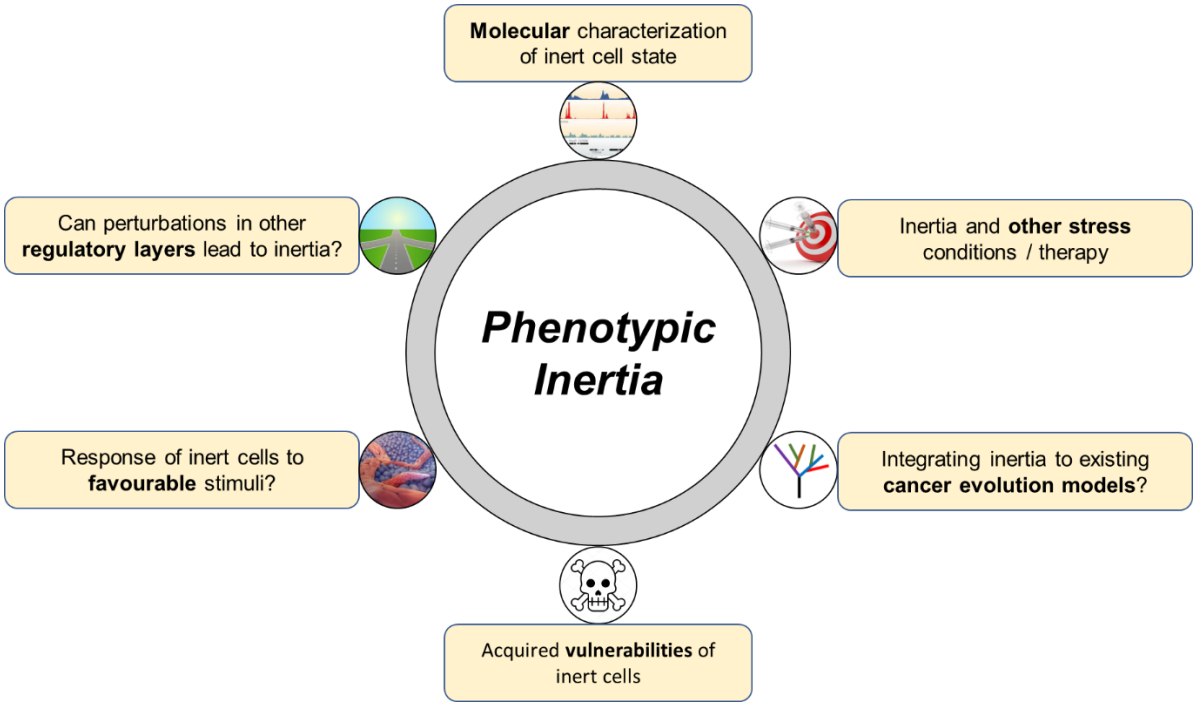


Figure 68: Follow-up questions on causes and consequences of phenotypic inertia in cancer evolution raised by my findings

5.6 Concluding remarks

I herein propose a network-level effect, where inactivation of multiple diverse epigenetic regulators results in the inability of cells to efficiently rewire their transcriptional response to stress. This leads to phenotypic unresponsiveness (Inertia), making the cells less likely to die upon exposure to stress, thus increasing the probability of acquiring secondary adaptive traits to ensure long-term survival [Figure 69].

The novelty of my PhD thesis resides in the experimental approach that consists of systematic functional perturbation of the entire epigenetic network to uncover potential fitness relationships between the examined properties of the system. On top of that phenotypes like plasticity (i.e. state transitions) are not inferred by the data but are rather directly monitored. The above are in contrast with most of the existing literature that consists of correlative explorations between epigenetic properties and transcriptional dynamics in cancer cells, thus limiting the power to draw causal links among them. Of note, the experimental procedures described in this thesis fall into the greater category of reductionism. This allows the elimination of varying confounding properties within the system to uncover potential patterns and functional links. Subsequent extrapolation to the more complex *in vivo* setting needs caution and requires further examination as described in greater detail in 5.2.3.

The observations made in this thesis may have significant implications for cancer biology. First, the widespread selective advantage of epigenetically deregulated cells under diverse hostile conditions provides a potential explanation for the recurrent inactivation of this regulatory layer at the later stages of tumorigenesis. Second, it is a proof of principle that fitness relationships upon perturbation should be examined under various contexts, including environmental challenges. Third, the fact that multiple data from the fitness assays and transcriptomic analysis could be explained by conceptually antithetic models like cellular plasticity and phenotypic inertia, reinforce the need to utilize better experimental frameworks to assign observations to models underlying cancer evolution. Finally, the fact that targeted therapy can induce similar transcriptional changes in cancer cells, along the stress-fitness axis that was explored in this study, points to a potential relevance of phenotypic inertia in therapy resistance and raises caution on the use of targeted therapy against epigenetic regulators in the clinical setting.

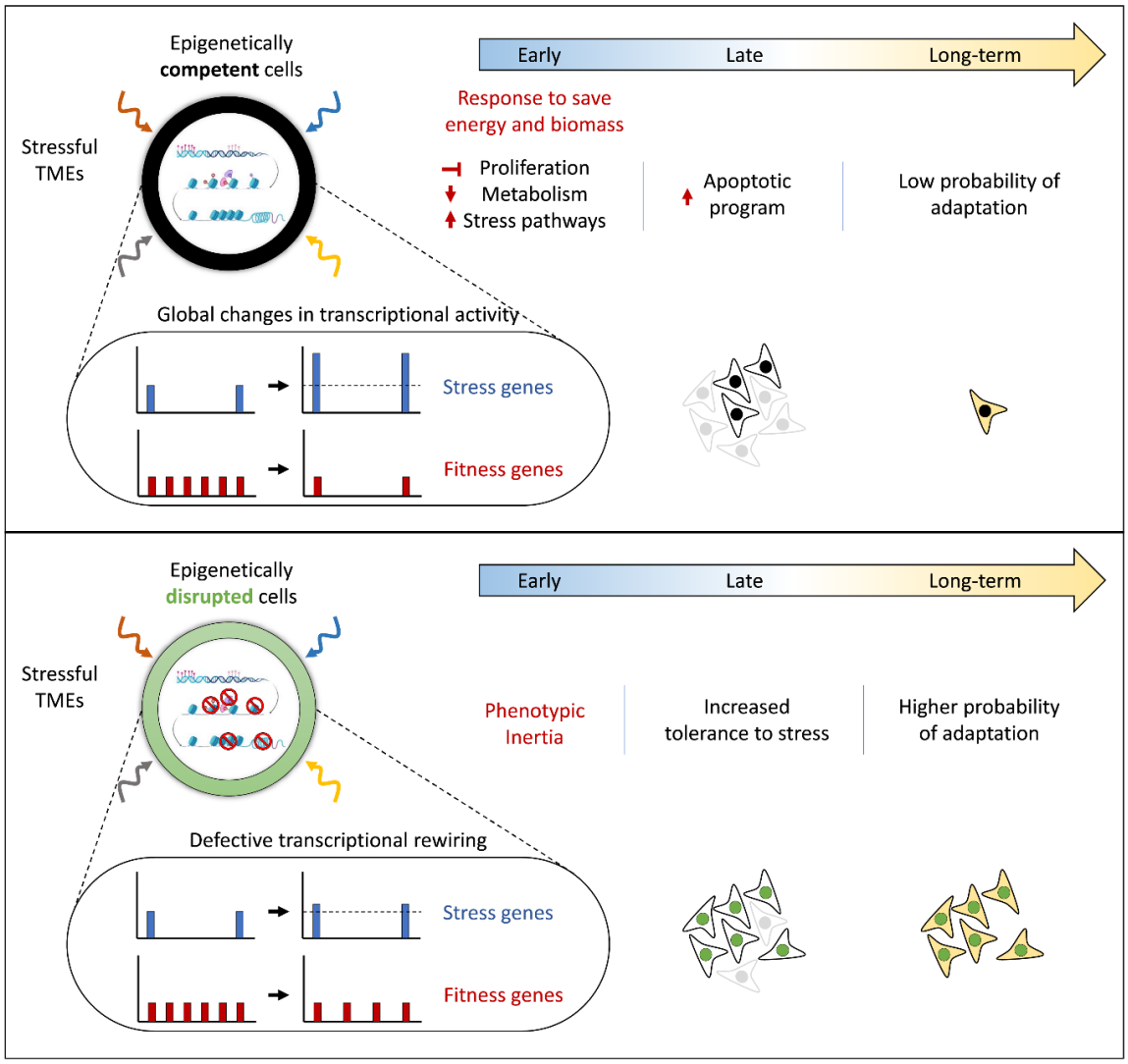


Figure 69: Proposed model of phenotypic inertia upon epigenetic deregulation.

Tables

Table 1. Epigenetic regulators targeted in the large-scale fitness assays

ACTL6A	BRWD3	DOT1L	HIF1AN	KDM1A	MSL1	PRMT1	SIRT3	TET2
ALKBH1	CARM1	EED	HIRA	KDM2A	MSL3	PRMT2	SIRT4	TET3
ARID1A	CBX1	EHMT1	HIRIP3	KDM2B	MTA1	PRMT3	SIRT5	TRIM24
ARID1B	CBX2	EHMT2	HIST1H1A	KDM3A	MTA2	PRMT5	SIRT6	TRIM28
ARID2	CBX3	ELP3	HIST1H1B	KDM3B	MTA3	PRMT6	SIRT7	TRIM66
ARID3A	CBX4	EP300	HIST1H1C	KDM4A	NAP1L1	PRMT7	SMARCA1	UBE2A
ARID3B	CBX5	EP400	HIST1H1D	KDM4B	NAP1L3	PRMT8	SMARCA2	UBE2B
ARID4A	CBX6	ERCC6	HIST1H1E	KDM4C	NAP1L4	PRMT9	SMARCA4	UHRF1
ARID4B	CBX7	EZH1	HIST1H1T	KDM4D	NAP1L5	RBBP4	SMARCA5	UHRF2
ARID5A	CBX8	EZH2	HIST2H3A	KDM4E	NCOA1	RBBP7	SMARCAD1	USP3
ARID5B	CDY2A	GFI1B	HIST3H3	KDM5A	NCOA3	RBM10	SMARCAL1	UTY
ASF1A	CDYL	GTF3C4	HJURP	KDM5B	NCOR1	RCC1	SMARCB1	YEATS4
ASF1B	CDYL2	H1F0	HLTF	KDM5C	NEDD4	RCOR1	SMARCC1	ZBTB33
ASH1L	CECR2	H1FNT	HMGA1	KDM5D	NPM2	RCOR2	SMARCC2	ZBTB4
ATAD2	CENPA	H1FOO	HMGA2	KDM6A	NR3C1	RING1	SMARCD1	ZBTB7C
ATAD2B	CENPE	H1FX	HMGB1	KDM6B	NSD1	RIOX1	SMARCD2	ZMYND11
ATF2	CENPF	H2AFB1	HMGB2	KDM7A	NSD2	RIOX2	SMARCD3	ZMYND8
ATRX	CHAF1A	H2AFJ	HMGB3	KMT2A	NSD3	RNF2	SMARCE1	TNP1
BAP1	CHAF1B	H2AFV	HMGN1	KMT2B	NUP153	RNF20	SMCHD1	TNP2
BAZ1A	CHD1	H2AFX	HMGN2	KMT2C	NUP62	RNF40	SMYD1	HMGB4
BAZ1B	CHD1L	H2AFY	HMGN3	KMT2D	PADI4	RRP8	SMYD2	SMC1B
BAZ2A	CHD2	H2AFY2	HMGN5	KMT2E	PBRM1	RSF1	SMYD3	DNMT3L
BAZ2B	CHD3	H2AFZ	HP1BP3	KMT5A	PCGF1	RUVBL1	SMYD5	
BCOR	CHD4	H3F3B	HR	KMT5B	PCGF2	SAFB	SP100	
BMI1	CHD5	HAT1	HSPBAP1	KMT5C	PCGF3	SATB1	SRCAP	
BPTF	CHD6	HDAC1	IKZF1	L3MBTL1	PCGF5	SATB2	SS18L1	
BRD1	CHD7	HDAC10	INCENP	MAF	PCGF6	SCMH1	SSRP1	
BRD2	CHD8	HDAC11	INO80	MBD1	PHC1	SETD1A	STAG1	
BRD3	CHD9	HDAC2	JARID2	MBD2	PHC2	SETD1B	SUDS3	
BRD4	CHRAC1	HDAC3	JMJD4	MBD3	PHC3	SETD2	SUV39H1	
BRD7	CREBBP	HDAC4	KAT2A	MBD3L1	PHF1	SETD6	SUV39H2	
BRD8	CTCF	HDAC5	KAT2B	MBD4	PHF2	SETD7	SUZ12	
BRD9	DDX11	HDAC6	KAT5	MBD5	PHF8	SETDB1	TADA2A	
BRDT	DMAP1	HDAC7	KAT6A	MBD6	POLE3	SETDB2	TAF1	
BRPF1	DNMT1	HDAC8	KAT6B	MECP2	PRDM2	SETMAR	TAF5	
BRPF3	DNMT3A	HDAC9	KAT7	MEN1	PRDM8	SIRT1	TAF6L	
BRWD1	DNMT3B	HELLS	KAT8	MPHOSPH8	PRDM9	SIRT2	TET1	

Non-expressed genes that were used as negative controls throughout this study are depicted in grey.

Table 2. Cellular models used in this study

Cell line name	Application	Culture conditions
PDX L1C5c	Parental clonal lines; Exome-seq	RPMI 1640 supplemented with 10% FBS, 100 U/ml penicillin, and 100 µg/ml streptomycin at 37 °C in 5% C02
PDX MeA5a		
PDX MeA5a eGFP – NLS	*In vitro* competition assay	
PDX MeA5a mCheryy – NLS		
PDX MeA5a Glucose FRET sensor	Live cell imaging	
HEK293T	Virus generation	
PDX MeA5a KO library	Large-scale fitness assays and subsequent experiments	
PDX L1C5c KO library		
PDX MeA5a KO sgRNA barcoded	*In vivo* competition assay; Bulk RNA-seq; scRNA-seq;	

Table 3. Characteristics of Patient Derived Xenograft (PDX) models

		PDX MEXF 2090	PDX LXFL 1674
Patient information	Gender	female	female
	Age at surgery	68	45
Patient tumor information	Histology	Melanoma	Large Cell Carcinoma
	Stage at surgery	Not available	Not available
	Differentiation	Good	Poor
	Therapy prior to surgery	Not known	Not known
	Origin of xenograft	Primary (Skin)	Primary (Lung)
PDX histological characteristics	Stroma content	3.00%	15.00%
	Vascularization	Low	Intermediate
	Grading	NA	Undifferentiated

Table 4. Plasmids used in this study

Plasmid name	Use	Growth conditions	Source
pCW-Cas9	Packaging of lentiviral particles	Stbl3 / 37 °C / Amp	Addgene; #50661
psPax2		DH5a / 37 °C / Amp	Addgene; #12260
pMD2.G		DH5a / 37 °C / Amp	Addgene; #12259
pAdVAntage™ Vector		Stbl3 / 37 °C / Amp	Promega; #E1711

Arrayed lentiviral sgRNA library	KO induction in epigenetic regulators	NA	(Henser-Brownhill, Monserrat & Scaffidi, 2017)
pUAS-mCherry-NLS	Cloning	DH5a / 37 °C / Amp	Addgene; #87695
pTRIP-SFFV-mCherry-NLS	Generation of fluorescently labelled melanoma cells	NEB Stable / 30 °C / Amp	Generated in this study
pTRIP-SFFV-EGFP-NLS			Addgene; #86677
PiggyBac transposon carrying FRET biosensor	Generation of melanoma cells expressing a glucose FRET sensor	Stbl3 / 37 °C / Amp	Eric Sahai; (Kondo et al., 2021)
PB transposase (PBase)		Stbl3 / 37 °C / Amp	Eric Sahai; (Liang et al., 2009)
PUC57-Capture_sequence_1	Cloning	DH5a / 37 °C / Amp	Genscript
pLenti-BSD-sgRNA-Capture_seq_1E	Barcoding KO populations for multiplexed scRNA-seq	ccdB Survival/ 37°C / Amp + Chl	Generated in this study

Table 5. Primers and other oligonucleotides used in this study

Primers used for qRT-PCR to assess total mRNA levels

Primer name	Sequence (5' to 3')
IL1B Forward	GCTTATTACAGTGGCAATGAGG
IL1B Reverse	AGATTCGTAGCTGGATGCC
SAT1 Forward	GAAGAATCTAAGCCAGGTTGC
SAT1 Reverse	ATGGATGGTTCATTCCATTCTG
MTIE Forward	CAAGAAGAGCTGCTGTTCC
MTIE Reverse	AGAGCTGTTCCCACATCAG
GLRX Forward	GTCTCTTTGCAACAGAGTGG
GLRX Reverse	TTCCTATGAGATCTGTGGTTACTG
TIMM13 Forward	AATGGAGCAGAGGATGACG
TIMM13 Reverse	ACTTCTGCTCGGAGTTGTC
SNRPA1 Forward	CACCAATAATAGTCTCGTGGA
SNRPA1 Reverse	CTTAGGATACTTAGGTAAGTCAGC
DDX21 Forward	TGCCATCCCTTTGATTGAG
DDX21 Reverse	GCAAGAACCAGTACCTGAG
PPIA Forward	AGACTGAGTGGTTGGATGG
PPIA Reverse	ATCTTCTTGCTGGTCTTGC

Primers used to replace GFP with mCherry in pTRIP-SFFV-EGFP-NLS

Primer name	Sequence (5' to 3')
mCherry_NLS amplification Forward	TAGCGGATCCATGGTGAGCAAGGGCGAG
mCherry_NLS amplification Reverse	AGGTCTCGAGTCTTATCATGTCTGCTCGAAGC

Primers used for library preparation and NGS for the in vivo competition experiment

Primer name	Sequence (5' to 3')
sgRNA outter PCR1 Forward	AGAGGTACCAAGGTCTGGCA
sgRNA outter PCR1 Reverse	CCTCGACCTGCTGGAATCTC
Miseq inner PCR2 Forward	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGTTTCTTGG GTAGTTTGCAGTTTT
Miseq inner PCR2 Reverse	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGCACCGA CTCGGTGCCACTTTT

Synthetic construct carrying an engineered sgRNA scaffold with a 3' terminal Capture Sequence 1 that is compatible with the 10x Genomics Feature barcoding technology. Restriction sites for BstBI and NsiI are labelled in blue.

Primer name	Sequence (5' to 3')
Synthetic sgRNA scaffold with Capture Sequence 1	**TTCGAA**CGTCTCTGTTTAAGAGCTAAGCTGGAAACAGCATAGCAAGTT TAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGG TGC**GCTTTAAGGCCGGTCCTAGCAA**TTTTTTTCTCGAGGTCGACGGTA TCGATAAGCTCGCTTCACGAGATTCCAGCAGGTCGAGGGACCTAATAA CTTCGTATAGCATACATTATACGAAGTTATATTAAGGGTTCCAAGCTTA AGCGGCCGCGTGGATAACCGTATTACCGCC**ATGCAT**

Sequences containing the sgRNAs against the indicated genes along with a fixed 10X Genomic barcode, used for population calling in the single cell transcriptomics experiment

sgRNA target	Sequence (5' to 3')	Sequence (5' to 3')
EED	GGCGTGTTTGTAGGTGTATC	(BC)GTTTAAGAGCTAAGCTGGAA
EZH2	ACACGCTTCCGCCAACAAAC	(BC)GTTTAAGAGCTAAGCTGGAA
HIST1H1B	CGCTTTGCGCTTAGCAGCGC	(BC)GTTTAAGAGCTAAGCTGGAA
SMARCD1	GAAACGGCTAGATATCCAAG	(BC)GTTTAAGAGCTAAGCTGGAA
SUZ12	ACGGCTTCGGGCGGCAAATC	(BC)GTTTAAGAGCTAAGCTGGAA
TNP2	TCTGGCTCCGGCTGCCACGA	(BC)GTTTAAGAGCTAAGCTGGAA

Primers used for Sanger sequencing

Primer name	Sequence (5' to 3')	
Illumina outer 1R	CCTCGACCTGCTGGAATCTC	Sequencing of pLenti plasmids to verify cloning of new sgRNA scaffold
Universal U6 promoter	GACTATCATATGCTTACCGT	Sequencing of pLenti plasmids to verify sgRNA integration

Table 6. Chemical compounds used in this study

Name	Target	Source	Catalogue number	Working concentration
RGFP966	HDAC3	APExBIO	#A8803	5 μM
Tubastatin A	HDAC6	APExBIO	#A4101	2.5 μM
EX 527 (SEN0014196)	SIRT1	APExBIO	#A4181	2.5 μM
EPZ004777	DOT1L	APExBIO	#A4170	2.5 μM
MM-102	KMT2A	APExBIO	#B1582	2.5 μM
WM-8014	KAT6A	APExBIO	#A8779	5 μM
GSK126	EZH2	Stratech	#S7061-SEL	5 μM
Quisinostat	pan HDAC	Insight Biotechnology	#HY-15433-1ml	5 nM
JIB-04	pan-Jumonji histone demethylase	APExBIO	#B1579	5 nM

Table 7. Antibodies used in this study

Abbreviations; IF: Immunofluorescence, IHC: Immunohistochemistry, IP: Intraperitoneal

Name	Application (Dilution)	Source	Catalogue number
Mouse monoclonal anti-Lamin A/C (636)	IF (1:200)	Santa Cruz Biotechnology	#sc-7292
Rabbit polyclonal anti-trimethyl-Histone H3 (Lys27)	IF (1:400) IHC (1:200)	EMD Millipore	#07-449
Rabbit polyclonal anti-acetyl-Histone H4 (Lys12)	IF (1:500)	Cell Signaling Technology	#2591
Bevacizumab (anti-VEGF recombinant antibody)	IP injection (2 mg/kg and 8 mg/kg)	Stratech	#A2006-SEL-5mg

Table 8: Stress-specific phenotype (z-score) of the indicated KO populations as determined in the large-scale fitness assays

Gene	Functional class	MEXF 2090			LXFL 1674
		Acidic conditions	Nutrient Starvation	Replicative stress	Nutrient Starvation
ACTL6A	Chromatin remodeller	1.75	-1.16	0.90	1.68
ALKBH1	DNA Me Eraser	-0.09	1.81	0.25	2.08
ARID1A	Chromatin remodeller	2.02	0.20	-0.31	2.62
ARID1B	Chromatin remodeller	1.97	0.61	-1.87	1.28
ARID2	Chromatin remodeller	2.14	-2.18	15.17	1.81
ARID3A	Chromatin remodeller	-0.46	-0.67	-0.25	-0.47
ARID3B	Chromatin remodeller	0.63	-0.18	0.78	-0.95
ARID4A	Chromatin remodeller	0.12	-0.46	0.18	-1.69
ARID4B	Chromatin remodeller	0.03	-3.40	-0.86	-2.13
ARID5A	Chromatin remodeller	2.64	1.14	0.12	3.96
ARID5B	Chromatin remodeller	-2.15	-2.53	0.69	-1.63
ASF1A	Histone variants and chaperons	-0.03	3.19	-0.07	-0.32
ASF1B	Histone variants and chaperons	-0.53	0.93	0.43	2.54
ASH1L	Histone Me Writer	3.03	1.45	-0.56	-0.61
ATAD2	Histone Ac Reader	-0.52	-1.95	-1.85	1.71
ATAD2B	Histone Ac Reader	0.81	1.61	0.84	0.59
ATF2	Histone Ac Writer	5.14	3.80	-0.32	3.06
ATRX	Chromatin remodeller	1.75	1.45	0.16	0.01
BAP1	Histone Ub Eraser	-0.66	-2.66	-0.88	0.61
BAZ1A	Chromatin remodeller	0.27	6.44	-0.05	-1.02
BAZ1B	Histone P Writer	8.19	7.57	1.17	-0.28
BAZ2A	Histone Ac Reader	0.14	-1.52	-1.46	1.79
BAZ2B	Histone Ac Reader	1.40	2.19	0.51	1.99
BCOR	Histone Ac Eraser	0.67	1.44	1.38	-0.34
BMI1	Histone Ub Writer	1.47	-0.08	-0.54	-0.08
BPTF	Histone Me Reader	0.80	0.76	1.89	-0.97
BRD1	Histone Ac Reader	1.99	0.89	0.17	-1.68
BRD2	Histone Ac Reader	0.48	-5.75	3.79	Failed QC
BRD3	Histone Ac Reader	0.20	-0.72	2.01	-1.54
BRD4	Histone Ac Reader	-0.20	-5.47	6.87	Failed QC
BRD7	Histone Ac Reader	Failed QC			0.01
BRD8	Histone Ac Reader	-0.49	-4.63	4.90	-1.05
BRD9	Histone Ac Reader	0.25	2.00	1.18	-0.85
BRDT	Histone Ac Reader	Not expressed			Not expressed
BRPF1	Histone Ac Reader	-0.29	-3.23	-0.59	-1.04
BRPF3	Histone Ac Reader	0.77	1.44	0.12	0.93
BRWD1	Histone Ac Reader	0.24	0.26	1.36	0.99
BRWD3	Rest	3.66	2.60	-1.44	1.54
CARM1	Histone Me Writer	3.30	-1.63	-0.65	-1.63
CBX1	Histone Me Reader	0.57	-1.44	1.02	1.34
CBX2	Histone Me Reader	0.31	0.40	1.04	-0.47
CBX3	Histone Me Reader	1.18	-2.57	-0.84	-0.31
CBX4	Histone Me Reader	0.78	-2.18	0.78	0.14
CBX5	Histone Me Reader	-0.76	-3.03	-0.14	-0.17
CBX6	Histone Me Reader	1.42	0.01	0.02	-0.57
CBX7	Histone Me Reader	Not expressed			Not expressed
CBX8	Histone Me Reader	0.46	-0.35	-0.33	0.79
CDY2A	Histone Ac Writer	Not expressed			Not expressed
CDYL	Histone Me Writer	0.24	10.36	-0.80	-0.78
CDYL2	Histone Me Reader	0.32	2.06	0.45	0.07
CECR2	Chromatin remodeller	Not expressed			Not expressed
CENPA	Histone variants and chaperons	Failed QC			Failed QC
CENPE	Rest	2.20	0.26	-0.86	Failed QC
CENPF	Rest	0.29	0.89	0.47	-1.37
CHAF1A	Histone variants and chaperons	1.24	1.63	2.56	3.57
CHAF1B	Histone variants and chaperons	3.71	4.39	2.10	Failed QC
CHD1	Chromatin remodeller	3.69	-0.21	-0.58	1.27
CHD1L	Chromatin remodeller	1.99	1.18	0.21	0.71
CHD2	Chromatin remodeller	3.58	1.49	0.01	0.40
CHD3	Chromatin remodeller	0.05	-0.29	0.31	-1.39
CHD4	Chromatin remodeller	-0.32	0.47	0.67	1.01
CHD5	Chromatin remodeller	Not expressed			Not expressed
CHD6	Chromatin remodeller	0.62	0.08	-1.12	-1.51
CHD7	Chromatin remodeller	3.19	1.00	1.39	1.85
CHD8	Chromatin remodeller	3.36	4.70	-0.18	0.08
CHD9	Chromatin remodeller	1.99	-2.98	-2.12	-0.75
CHRAC1	Chromatin remodeller	1.58	7.89	0.13	0.88
CREBBP	Histone Ac Writer	-2.58	0.46	4.02	4.50
CTCF	Genome topology	1.09	-0.68	2.08	-0.17
DDX11	Rest	0.71	1.86	1.43	Failed QC
DMAP1	Histone Ac Eraser	0.52	-0.42	1.19	-0.44
DNMT1	DNA Me Writer	0.27	-0.63	1.46	Failed QC
DNMT3A	DNA Me Writer	1.45	-1.23	2.15	1.36
DNMT3B	DNA Me Writer	1.72	-0.34	1.87	-0.32
DOT1L	Histone Me Writer	0.16	2.92	-1.11	-0.89
EED	Histone Me Writer	-2.00	29.90	-3.44	10.63
EHMT1	Histone Me Writer	2.18	5.60	-0.03	-1.56
EHMT2	Histone Me Writer	2.02	2.68	-1.23	-2.09
ELP3	Histone Ac Writer	-0.52	-0.57	2.19	-0.09
EP300	Histone Ac Writer	Failed QC			0.50
EP400	Histone Ac Writer	-2.62	-2.48	0.60	0.15
ERCC6	Chromatin remodeller	2.51	1.61	1.40	Not expressed
EZH1	Histone Me Writer	1.44	1.09	1.72	0.53
EZH2	Histone Me Writer	-0.68	31.61	0.05	7.37
GFI1B	Histone Ac Eraser	Not expressed			Not expressed
GTF3C4	Histone Ac Writer	-1.46	-1.20	2.11	Failed QC
H1F0	Histone variants and chaperons	0.61	0.65	-0.50	-1.49
H1FNT	Histone variants and chaperons	Not expressed			Not expressed
H1FOO	Histone variants and chaperons	Not expressed			Not expressed
H1FX	Histone variants and chaperons	0.26	-0.72	-0.47	-0.84
H2AFB1	Histone variants and chaperons	Not expressed			Not expressed
H2AFJ	Histone variants and chaperons	0.57	0.53	-0.92	-0.70
H2AFV	Histone variants and chaperons	5.46	4.52	1.76	-1.12
H2AFX	Histone variants and chaperons	Failed QC			-1.20
H2AFY	Histone variants and chaperons	0.47	0.58	-1.28	-1.14
H2AFY2	Histone variants and chaperons	Not expressed			-0.20
H2AFZ	Histone variants and chaperons	Failed QC			-1.94
H3F3B	Histone variants and chaperons	2.39	3.92	0.45	1.34
HAT1	Histone Ac Writer	3.05	-1.18	1.98	1.13
HDAC1	Histone Ac Eraser	0.15	0.27	0.00	2.07
HDAC10	Histone Ac Eraser	Not expressed			-1.44
HDAC11	Histone Ac Eraser	-0.22	0.37	-1.19	0.20

HDAC2	Histone Ac Eraser	0.60	-0.59	1.71	-0.19
HDAC3	Histone Ac Eraser	0.70	2.38	1.96	Failed QC
HDAC4	Histone Ac Eraser	0.28	0.16	-1.87	0.69
HDAC5	Histone Ac Eraser	0.46	1.24	-0.51	-1.48
HDAC6	Histone Ac Eraser	2.57	2.80	0.28	0.86
HDAC7	Histone Ac Eraser	0.53	2.30	2.65	1.29
HDAC8	Histone Ac Eraser	2.37	2.03	-1.57	-0.55
HDAC9	Histone Ac Eraser	0.48	1.29	0.22	Not expressed
HELLS	DNA Me Writer	2.09	6.52	-0.90	-0.87
HIF1AN	Histone Ac Eraser	0.49	2.89	2.40	0.34
HIRA	Histone variants and chaperons	2.18	3.46	-1.02	-0.12
HIRIP3	Histone variants and chaperons	0.13	-0.42	-0.19	1.74
HIST1H1A	Histone variants and chaperons	2.13	2.92	1.12	Not expressed
HIST1H1B	Histone variants and chaperons	3.30	7.62	-0.68	Not expressed
HIST1H1C	Histone variants and chaperons	3.08	2.60	-1.01	-1.23
HIST1H1D	Histone variants and chaperons	3.88	6.73	-1.54	Not expressed
HIST1H1E	Histone variants and chaperons	Not expressed			-0.35
HIST1H1T	Histone variants and chaperons	Not expressed			Not expressed
HIST2H3A	Histone variants and chaperons	Not expressed			Not expressed
HIST3H3	Histone variants and chaperons	Not expressed			Not expressed
HJURP	Histone variants and chaperons	Failed QC			Failed QC
HLTF	Chromatin remodeller	2.46	2.17	0.96	-0.52
HMGA1	Genome topology	2.05	1.85	-2.41	Failed QC
HMGA2	Genome topology	3.57	2.45	-1.19	-0.06
HMGB1	Genome topology	2.33	0.94	-3.92	Failed QC
HMGB2	Genome topology	4.72	3.12	1.33	3.06
HMGB3	Genome topology	3.99	1.44	0.51	3.52
HMGN1	Genome topology	5.51	3.45	2.59	1.64
HMGN2	Genome topology	2.11	1.60	-0.08	0.63
HMGN3	Genome topology	6.71	6.95	1.84	-0.30
HMGN5	Genome topology	4.74	3.87	2.42	-1.63
HP1BP3	Rest	4.25	2.85	1.59	1.83
HR	Histone Me Eraser	Not expressed			Not expressed
HSPBAP1	Rest	Not expressed			0.59
IKZF1	Histone Ac Eraser	Not expressed			Not expressed
INCENP	Rest	3.55	2.46	0.34	Failed QC
INO80	Chromatin remodeller	1.95	0.61	-1.63	Failed QC
JARID2	Histone Me Writer	2.16	0.85	2.09	0.52
JMJD4	Histone Me Eraser	0.82	0.87	-0.49	-0.84
KAT2A	Histone Ac Writer	3.93	9.38	1.04	1.72
KAT2B	Histone Ac Writer	3.52	2.72	1.70	0.36
KAT5	Histone Ac Writer	0.50	-2.04	4.69	1.52
KAT6A	Histone Ac Writer	2.04	4.96	1.49	-1.82
KAT6B	Histone Ac Writer	3.95	4.86	3.37	-1.41
KAT7	Histone Ac Writer	2.73	1.78	-0.53	-0.49
KAT8	Histone Ac Writer	Failed QC			Failed QC
KDM1A	Histone Me Eraser	-1.10	-2.22	-2.26	-2.45
KDM2A	Histone Me Eraser	-0.26	-2.15	0.87	-1.97
KDM2B	Histone Me Eraser	-0.08	-2.60	3.05	0.06
KDM3A	Histone Me Eraser	0.14	1.42	0.41	0.30
KDM3B	Histone Me Eraser	-1.19	4.70	-2.15	-1.06
KDM4A	Histone Me Eraser	-0.33	-0.21	0.52	1.01
KDM4B	Histone Me Eraser	-0.22	3.23	1.86	1.49
KDM4C	Histone Me Eraser	0.20	2.48	0.05	1.85
KDM4D	Histone Me Eraser	-0.27	2.14	-0.14	Not expressed
KDM4E	Histone Me Eraser	Not expressed			Not expressed
KDM5A	Histone Me Eraser	2.05	3.17	-0.35	-0.41
KDM5B	Histone Me Eraser	-0.17	0.61	-0.32	0.83
KDM5C	Histone Me Eraser	0.53	1.49	1.43	-1.95
KDM5D	Histone Me Eraser	Not expressed			Not expressed
KDM6A	Histone Me Eraser	-0.20	-0.87	0.46	-0.92
KDM6B	Histone Me Eraser	0.92	0.37	0.44	Failed QC
KDM7A	Histone Me Eraser	0.33	4.41	-1.43	0.45
KMT2A	Histone Me Writer	-0.87	2.38	-0.71	-1.44
KMT2B	Histone Me Writer	-0.34	0.42	-0.17	0.21
KMT2C	Histone Me Writer	-0.28	-0.60	-1.27	1.07
KMT2D	Histone Me Writer	-0.92	-0.25	0.90	0.17
KMT2E	Histone Me Writer	1.43	-0.86	1.15	1.05
KMT5A	Histone Me Writer	-0.87	-1.09	-0.36	1.64
KMT5B	Histone Me Writer	-0.55	-0.60	-0.40	0.04
KMT5C	Histone Me Writer	1.90	0.47	-0.59	0.82
L3MBTL1	Histone Me Reader	Not expressed			Not expressed
MAF	Histone Ac Writer	Not expressed			1.62
MBD1	DNA Me Reader	3.33	-0.06	-1.18	1.55
MBD2	DNA Me Reader	4.44	-0.85	0.38	1.59
MBD3	DNA Me Reader	5.26	7.65	3.42	6.12
MBD3L1	Histone Ac Eraser	Not expressed			Not expressed
MBD4	DNA Me Reader	1.76	0.18	3.65	1.35
MBD5	Histone Ub Eraser	2.57	-1.02	-0.18	Not expressed
MBD6	Histone Ub Eraser	7.10	-0.38	0.01	3.43
MECP2	DNA Me Reader	1.51	1.62	1.51	0.62
MEN1	Histone Me Writer	-0.91	-2.07	-2.75	2.32
MPHOSPH8	Histone Me Reader	-0.63	-0.82	-2.06	-1.29
MSL1	Histone Ac Writer	2.70	-1.73	0.96	-0.84
MSL3	Histone Ac Writer	0.95	-1.79	-0.37	-1.19
MTA1	Chromatin remodeller	6.19	6.26	1.43	0.71
MTA2	Chromatin remodeller	7.07	6.95	5.01	-3.06
MTA3	Chromatin remodeller	2.61	2.16	0.88	-0.18
NAP1L1	Histone variants and chaperons	0.36	-0.42	-0.30	0.72
NAP1L3	Histone variants and chaperons	Not expressed			Not expressed
NAP1L4	Histone variants and chaperons	2.16	2.80	-0.14	-0.54
NAP1L5	Histone variants and chaperons	1.10	3.72	-0.78	Not expressed
NCOA1	Histone Ac Writer	4.96	2.56	-0.18	0.94
NCOA3	Histone Ac Writer	1.34	0.83	0.06	-1.14
NCOR1	Histone Ac Eraser	2.47	9.61	6.80	-1.22
NEDD4	Histone Ub Writer	1.56	2.61	-0.72	-0.94
NPM2	Histone variants and chaperons	Not expressed			-1.86
NR3C1	Rest	3.91	2.56	1.12	-0.54
NSD1	Histone Me Writer	3.89	-2.18	2.73	-2.34
NSD2	Histone Me Writer	0.68	-0.03	0.28	-0.47
NSD3	Histone Me Writer	2.09	0.75	0.57	-1.31
NUP153	Rest	2.66	-0.26	-0.23	Failed QC
NUP62	Rest	2.93	-0.77	1.04	1.87
PADI4	Rest	Not expressed			Not expressed
PBRM1	Chromatin remodeller	3.61	-2.21	7.65	0.95
PCGF1	Histone Ub Writer	-2.62	0.00	2.82	0.55
PCGF2	Histone Ub Writer	0.65	1.09	0.51	0.07
PCGF3	Histone Ub Writer	-0.51	2.98	1.14	0.93
PCGF5	Histone Ub Writer	0.86	1.89	-0.02	-1.23
PCGF6	Histone Ub Writer	-0.66	-1.78	0.76	-1.56
PHC1	Histone Ub Writer	1.19	0.72	-0.46	0.68

PHC2	Histone Ub Writer	1.39	1.21	0.29	0.83
PHC3	Histone Ub Writer	1.25	3.51	1.26	-0.07
PHF1	Histone Me Reader	2.09	1.21	0.97	-0.38
PHF2	Histone Me Reader	1.07	-1.84	0.36	-1.12
PHF8	Histone Me Reader	0.89	-0.28	0.12	-3.61
POLE3	Chromatin remodeller	6.87	7.13	0.40	-0.44
PRDM2	Histone Me Writer	0.59	0.10	0.14	0.12
PRDM8	Histone Me Writer	-0.89	0.37	0.49	1.15
PRDM9	Histone Me Writer	Not expressed			Not expressed
PRMT1	Histone Me Writer	2.20	0.73	0.07	Failed QC
PRMT2	Histone Me Writer	-0.35	-0.71	-0.61	-1.78
PRMT3	Histone Me Writer	4.64	4.00	0.56	0.29
PRMT5	Histone Me Writer	0.68	1.22	-0.64	Failed QC
PRMT6	Histone Me Writer	-0.38	-0.39	-0.72	-1.18
PRMT7	Histone Me Writer	Failed QC			-0.21
PRMT8	Histone Me Writer	Not expressed			Not expressed
PRMT9	Histone Me Writer	2.03	3.18	-2.69	-0.05
RBBP4	Rest	1.61	0.56	-0.51	Failed QC
RBBP7	Rest	Failed QC			-1.69
RBM10	Rest	-1.13	-1.91	2.10	-2.36
RCC1	Genome topology	0.32	0.12	1.95	Failed QC
RCOR1	Histone Ac Eraser	-1.06	1.12	-3.02	-1.67
RCOR2	Histone Ac Eraser	Not expressed			-0.15
RING1	Histone Ub Writer	-0.39	3.54	0.79	0.48
RIOX1	Histone Me Eraser	-1.83	1.38	-1.71	-0.83
RIOX2	Histone Me Eraser	-0.17	2.98	-1.53	1.64
RNF2	Histone Ub Writer	2.84	1.68	1.12	0.12
RNF20	Histone Ub Writer	0.58	1.57	0.25	0.03
RNF40	Histone Ub Writer	0.20	1.27	0.65	0.86
RRP8	Histone Me Reader	1.88	0.03	-0.73	-0.30
RSF1	Rest	4.67	6.03	1.02	0.30
RUVBL1	Histone Ac Writer	3.28	1.19	2.10	-0.11
SAFB	Genome topology	4.26	5.00	2.37	2.92
SATB1	Rest	Failed QC			Not expressed
SATB2	Histone Ac Eraser	0.63	-0.51	-0.79	2.69
SCMH1	Histone Ub Writer	0.92	1.02	-1.11	0.80
SETD1A	Histone Me Writer	Failed QC			Failed QC
SETD1B	Histone Me Writer	Failed QC			-0.32
SETD2	Histone Me Writer	Failed QC			-3.18
SETD6	Histone Me Writer	Not expressed			-2.26
SETD7	Histone Me Writer	0.08	-0.46	0.09	-0.66
SETDB1	Histone Me Writer	Failed QC			-4.80
SETDB2	Histone Me Writer	-0.05	0.07	-0.22	Not expressed
SETMAR	Histone Me Writer	0.13	-0.72	-0.66	-1.30
SIRT1	Histone Ac Eraser	2.05	4.08	-1.29	0.20
SIRT2	Histone Ac Eraser	-0.61	-0.12	-2.05	0.20
SIRT3	Histone Ac Eraser	-1.66	1.00	0.59	0.37
SIRT4	Histone Ac Eraser	Not expressed			Not expressed
SIRT5	Histone Ac Eraser	1.85	4.24	-1.74	0.18
SIRT6	Histone Ac Eraser	-0.57	-1.92	-0.63	-1.32
SIRT7	Histone Ac Eraser	0.15	0.03	2.68	-0.70
SMARCA1	Chromatin remodeller	1.62	1.71	1.40	0.21
SMARCA2	Chromatin remodeller	1.99	-0.01	1.02	0.86
SMARCA4	Chromatin remodeller	0.81	3.73	0.03	4.13
SMARCA5	Chromatin remodeller	1.70	0.14	0.43	0.72
SMARCAD1	Chromatin remodeller	4.82	2.98	2.47	3.15
SMARCAL1	Chromatin remodeller	0.94	1.95	-0.10	-0.13
SMARCB1	Chromatin remodeller	1.99	2.99	1.32	Failed QC
SMARCC1	Chromatin remodeller	1.82	5.10	1.11	2.06
SMARCC2	Chromatin remodeller	-0.13	0.14	0.18	3.02
SMARCD1	Chromatin remodeller	0.59	6.20	0.65	1.00
SMARCD2	Chromatin remodeller	-0.99	0.23	-0.58	-0.77
SMARCD3	Chromatin remodeller	0.39	-0.85	-0.95	2.07
SMARCE1	Chromatin remodeller	-0.27	-1.45	-0.22	3.56
SMCHD1	Genome topology	1.07	0.25	-0.06	0.28
SMYD1	Histone Me Writer	Not expressed			Not expressed
SMYD2	Histone Me Writer	0.71	-0.43	1.15	0.57
SMYD3	Histone Me Writer	3.89	2.63	0.63	-0.73
SMYD5	Histone Me Writer	2.42	2.80	-1.16	-0.25
SP100	Rest	0.67	4.64	0.05	0.34
SRCAP	Chromatin remodeller	-0.20	-1.20	0.67	-0.55
SS18L1	Chromatin remodeller	1.36	1.19	0.37	0.56
SSRP1	Chromatin remodeller	0.16	-1.94	1.96	1.95
STAG1	Genome topology	0.86	-0.89	-0.90	0.18
SUDS3	Histone Ac Eraser	-1.51	0.62	-2.36	Failed QC
SUV39H1	Histone Me Writer	0.61	0.29	-0.15	-0.04
SUV39H2	Histone Me Writer	0.87	-0.73	1.06	-0.20
SUZ12	Histone Me Writer	-1.43	20.95	-2.68	10.86
TADA2A	Histone Ac Writer	3.58	3.67	1.78	-0.05
TAF1	Histone Ac Writer	1.31	-0.23	2.04	-0.62
TAF5	Histone Ac Writer	-0.62	-1.62	2.76	Failed QC
TAF6L	Histone Ac Writer	Not expressed			0.16
TET1	DNA Me Eraser	-0.26	-0.98	0.14	0.13
TET2	DNA Me Eraser	0.60	-0.70	1.38	Not expressed
TET3	DNA Me Eraser	0.96	-0.95	-0.15	0.70
TRIM24	Rest	0.29	2.83	-0.46	0.97
TRIM28	Rest	Failed QC			-3.08
TRIM66	Histone Ac Eraser	-1.16	-1.51	0.28	Not expressed
UBE2A	Histone Ub Writer	0.42	1.86	-0.19	-0.84
UBE2B	Histone Ub Writer	1.07	0.85	1.42	-0.59
UHRF1	DNA Me Writer	Failed QC			-1.77
UHRF2	DNA Me Reader	1.47	1.71	-0.69	1.91
USP3	Histone Ub Eraser	3.62	1.63	3.04	-1.53
UTY	Histone Me Eraser	Not expressed			Not expressed
YEATS4	Histone Ac Writer	2.93	0.37	0.15	-0.37
ZBTB33	DNA Me Reader	1.44	1.00	0.69	-0.77
ZBTB4	DNA Me Reader	0.43	-1.17	3.09	2.10
ZBTB7C	Rest	Not expressed			Not expressed
ZMYND11	Histone Me Reader	-0.09	1.08	-0.59	-0.87
ZMYND8	Rest	0.51	1.29	3.16	-1.37

Bibliography

Aissa, A.F., Islam, A.B.M.M.K., Ariss, M.M., Go, C.C., Rader, A.E., Conrardy, R.D., Gajda, A.M., Rubio-Perez, C., Valyi-Nagy, K., Pasquinelli, M., Feldman, L.E., Green, S.J., Lopez-Bigas, N., Frolov, M.V. & Benevolenskaya, E.V. (2021) Single-cell transcriptional changes associated with drug tolerance and response to combination therapies in cancer. *Nature Communications*. 12 (1), 1628. doi:10.1038/s41467-021-21884-z.

Allis, C.D. & Jenuwein, T. (2016) The molecular hallmarks of epigenetic control. *Nature Reviews. Genetics*. 17 (8), 487–500. doi:10.1038/nrg.2016.59.

Almagro, J., Messal, H.A., Elosegui-Artola, A., van Rheenen, J. & Behrens, A. (2022) Tissue architecture in tumor initiation and progression. *Trends in Cancer*. 8 (6), 494–505. doi:10.1016/j.trecan.2022.02.007.

Almouzni, G. & Cedar, H. (2016) Maintenance of Epigenetic Information. *Cold Spring Harbor Perspectives in Biology*. 8 (5), a019372. doi:10.1101/cshperspect.a019372.

Aprile-Garcia, F., Tomar, P., Hummel, B., Khavaran, A. & Sawarkar, R. (2019) Nascent-protein ubiquitination is required for heat shock–induced gene downregulation in human cells. *Nature Structural & Molecular Biology*. 26 (2), 137–146. doi:10.1038/s41594-018-0182-x.

Baggiolini, A., Callahan, S.J., Montal, E., Weiss, J.M., Trieu, T., et al. (2021) Developmental chromatin programs determine oncogenic competence in melanoma. *Science (New York, N.Y.)*. 373 (6559), eabc1048. doi:10.1126/science.abc1048.

Bar-Even, A., Paulsson, J., Maheshri, N., Carmi, M., O'Shea, E., Pilpel, Y. & Barkai, N. (2006) Noise in protein expression scales with natural protein abundance. *Nature Genetics*. 38 (6), 636–643. doi:10.1038/ng1807.

Barkley, D., Moncada, R., Pour, M., Liberman, D.A., Dryg, I., Werba, G., Wang, W., Baron, M., Rao, A., Xia, B., França, G.S., Weil, A., Delair, D.F., Hajdu, C., Lund, A.W., Osman, I. & Yanai, I. (2022) Cancer cell states recur across tumor types and form specific interactions with the tumor microenvironment. *Nature Genetics*. 54 (8), 1192–1201. doi:10.1038/s41588-022-01141-9.

Barnum, K.J. & O'Connell, M.J. (2014) Cell cycle regulation by checkpoints. *Methods in Molecular Biology (Clifton, N.J.)*. 1170, 29–40. doi:10.1007/978-1-4939-0888-2_2.

Baron, M., Tagore, M., Hunter, M.V., Kim, I.S., Moncada, R., Yan, Y., Campbell, N.R., White, R.M. & Yanai, I. (2020) The Stress-Like Cancer Cell State Is a Consistent Component of Tumorigenesis. *Cell Systems*. 11 (5), 536-546.e7. doi:10.1016/j.cels.2020.08.018.

Baylin, S.B. & Jones, P.A. (2011) A decade of exploring the cancer epigenome - biological and translational implications. *Nature Reviews. Cancer*. 11 (10), 726–734. doi:10.1038/nrc3130.

Berger, S.L., Kouzarides, T., Shiekhattar, R. & Shilatifard, A. (2009) An operational definition of epigenetics. *Genes & Development*. 23 (7), 781–783. doi:10.1101/gad.1787609.

Bhandari, V., Hoey, C., Liu, L.Y., Lalonde, E., Ray, J., et al. (2019) Molecular landmarks of tumor hypoxia across cancer types. *Nature Genetics*. 51 (2), 308–318. doi:10.1038/s41588-018-0318-2.

Binnewies, M., Roberts, E.W., Kersten, K., Chan, V., Fearon, D.F., Merad, M., Coussens, L.M., Gabrilovich, D.I., Ostrand-Rosenberg, S., Hedrick, C.C., Vonderheide, R.H., Pittet, M.J., Jain, R.K., Zou, W., Howcroft, T.K., Woodhouse, E.C., Weinberg, R.A. & Krummel, M.F. (2018) Understanding the tumor immune microenvironment (TIME) for effective therapy. *Nature Medicine*. 24 (5), 541–550. doi:10.1038/s41591-018-0014-x.

Bott, A.J., Peng, I.-C., Fan, Y., Faubert, B., Zhao, L., et al. (2015) Oncogenic Myc Induces Expression of Glutamine Synthetase through Promoter Demethylation. *Cell Metabolism*. 22 (6), 1068–1077. doi:10.1016/j.cmet.2015.09.025.

Boumahdi, S. & de Sauvage, F.J. (2020) The great escape: tumour cell plasticity in resistance to targeted therapy. *Nature Reviews. Drug Discovery*. 19 (1), 39–56. doi:10.1038/s41573-019-0044-1.

Boveri, T. (2008) Concerning the origin of malignant tumours by Theodor Boveri. Translated and annotated by Henry Harris. *Journal of Cell Science*. 121 Suppl 1, 1–84. doi:10.1242/jcs.025742.

Brennan, C.W., Verhaak, R.G.W., McKenna, A., Campos, B., Noushmehr, H., et al. (2013) The somatic genomic landscape of glioblastoma. *Cell*. 155 (2), 462–477. doi:10.1016/j.cell.2013.09.034.

Brinkman, E.K., Kousholt, A.N., Harmsen, T., Leemans, C., Chen, T., Jonkers, J. & van Steensel, B. (2018) Easy quantification of template-directed CRISPR/Cas9 editing. *Nucleic Acids Research*. 46 (10), e58. doi:10.1093/nar/gky164.

Brock, A., Krause, S. & Ingber, D.E. (2015) Control of cancer formation by intrinsic genetic noise and microenvironmental cues. *Nature Reviews. Cancer*. 15 (8), 499–509. doi:10.1038/nrc3959.

Brocks, D., Assenov, Y., Minner, S., Bogatyrova, O., Simon, R., et al. (2014) Intratumor DNA methylation heterogeneity reflects clonal evolution in aggressive prostate cancer. *Cell Reports*. 8 (3), 798–806. doi:10.1016/j.celrep.2014.06.053.

Burrell, R.A., McGranahan, N., Bartek, J. & Swanton, C. (2013) The causes and consequences of genetic heterogeneity in cancer evolution. *Nature*. 501 (7467), 338–345. doi:10.1038/nature12625.

Cancer Genome Atlas Research Network, Ley, T.J., Miller, C., Ding, L., Raphael, B.J., et al. (2013) Genomic and epigenomic landscapes of adult de novo acute myeloid leukemia. *The New England Journal of Medicine*. 368 (22), 2059–2074. doi:10.1056/NEJMoa1301689.

Carlson, M. & Laurent, B.C. (1994) The SNF/SWI family of global transcriptional activators. *Current Opinion in Cell Biology*. 6 (3), 396–402. doi:10.1016/0955-0674(94)90032-9.

Cen, H., Mao, F., Aronchik, I., Fuentes, R.J. & Firestone, G.L. (2008) DEVD-NucView488: a novel class of enzyme substrates for real-time detection of caspase-3 activity in live cells. *The FASEB Journal*. 22 (7), 2243–2252. doi:10.1096/fj.07-099234.

Chaligne, R., Gaiti, F., Silverbush, D., Schiffman, J.S., Weisman, H.R., et al. (2021) Epigenetic encoding, heritability and plasticity of glioma transcriptional cell states. *Nature Genetics*. 53 (10), 1469–1479. doi:10.1038/s41588-021-00927-7.

Chen, C.-H., Xiao, T., Xu, H., Jiang, P., Meyer, C.A., Li, W., Brown, M. & Liu, X.S. (2018) Improved design and analysis of CRISPR knockout screens. *Bioinformatics (Oxford, England)*. 34 (23), 4095–4101. doi:10.1093/bioinformatics/bty450.

Chen, J.F. & Yan, Q. (2021) The roles of epigenetics in cancer progression and metastasis. *The Biochemical Journal*. 478 (17), 3373–3393. doi:10.1042/BCJ20210084.

Cingolani, P., Platts, A., Wang, L.L., Coon, M., Nguyen, T., Wang, L., Land, S.J., Lu, X. & Ruden, D.M. (2012) A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: *Fly*. 6 (2), 80–92. doi:10.4161/fly.19695.

Ciriello, G. & Magnani, L. (2021) The many faces of cancer evolution. *iScience*. 24 (5), 102403. doi:10.1016/j.isci.2021.102403.

Concepcion, C.P., Ma, S., LaFave, L.M., Bhutkar, A., Liu, M., et al. (2022) Smarca4 Inactivation Promotes Lineage-Specific Transformation and Early Metastatic Features in the Lung. *Cancer Discovery*. 12 (2), 562–585. doi:10.1158/2159-8290.CD-21-0248.

Corbet, C., Draoui, N., Polet, F., Pinto, A., Drozak, X., Riant, O. & Feron, O. (2014) The SIRT1/HIF2α axis drives reductive glutamine metabolism under chronic acidosis and alters tumor response to therapy. *Cancer Research*. 74 (19), 5507–5519. doi:10.1158/0008-5472.CAN-14-0705.

Corces, M.R., Buenrostro, J.D., Wu, B., Greenside, P.G., Chan, S.M., Koenig, J.L., Snyder, M.P., Pritchard, J.K., Kundaje, A., Greenleaf, W.J., Majeti, R. & Chang, H.Y. (2016) Lineage-specific and single-cell chromatin accessibility charts human hematopoiesis and leukemia evolution. *Nature Genetics*. 48 (10), 1193–1203. doi:10.1038/ng.3646.

Cugusi, S., Mitter, R., Kelly, G.P., Walker, J., Han, Z., Pisano, P., Wierer, M., Stewart, A. & Svejstrup, J.Q. (2022) Heat shock induces premature transcript termination and reconfigures the human transcriptome. *Molecular Cell*. 82 (8), 1573-1588.e10. doi:10.1016/j.molcel.2022.01.007.

Daigeler, A., Klein-Hitpass, L., Chromik, M.A., Müller, O., Hauser, J., Homann, H.-H., Steinau, H.-U. & Lehnhardt, M. (2008) Heterogeneous in vitro effects of doxorubicin on gene expression in primary human liposarcoma cultures. *BMC cancer*. 8, 313. doi:10.1186/1471-2407-8-313.

De Conti, G., Dias, M.H. & Bernards, R. (2021) Fighting Drug Resistance through the Targeting of Drug-Tolerant Persister Cells. *Cancers*. 13 (5), 1118. doi:10.3390/cancers13051118.

Deans, C. & Maggert, K.A. (2015) What do you mean, 'epigenetic'? *Genetics*. 199 (4), 887–896. doi:10.1534/genetics.114.173492.

Delhommeau, F., Dupont, S., Valle, V.D., James, C., Trannoy, S., et al. (2009) Mutation in TET2 in Myeloid Cancers. *New England Journal of Medicine*. 360 (22), 2289–2301. doi:10.1056/NEJMoa0810069.

Denny, S.K., Yang, D., Chuang, C.-H., Brady, J.J., Lim, J.S., Grüner, B.M., Chiou, S.-H., Schep, A.N., Baral, J., Hamard, C., Antoine, M., Wislez, M., Kong, C.S., Connolly, A.J., Park, K.-S., Sage, J., Greenleaf, W.J. & Winslow, M.M. (2016) Nfib Promotes Metastasis through a Widespread Increase in Chromatin Accessibility. *Cell*. 166 (2), 328–342. doi:10.1016/j.cell.2016.05.052.

Dentro, S.C., Leshchiner, I., Haase, K., Tarabichi, M., Wintersinger, J., et al. (2021) Characterizing genetic intra-tumor heterogeneity across 2,658 human cancer genomes. *Cell*. 184 (8), 2239-2254.e39. doi:10.1016/j.cell.2021.03.009.

Di Gregorio, A., Bowling, S. & Rodriguez, T.A. (2016) Cell Competition and Its Role in the Regulation of Cell Fitness from Development to Cancer. *Developmental Cell*. 38 (6), 621–634. doi:10.1016/j.devcel.2016.08.012.

Di Tommaso, P., Chatzou, M., Floden, E.W., Barja, P.P., Palumbo, E. & Notredame, C. (2017) Nextflow enables reproducible computational workflows. *Nature Biotechnology*. 35 (4), 316–319. doi:10.1038/nbt.3820.

Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M. & Gingeras, T.R. (2013) STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*. 29 (1), 15–21. doi:10.1093/bioinformatics/bts635.

Duan, Y., Tian, L., Gao, Q., Liang, L., Zhang, W., Yang, Y., Zheng, Y., Pan, E., Li, S. & Tang, N. (2016) Chromatin remodeling gene ARID2 targets cyclin D1 and cyclin E1 to suppress hepatoma cell progression. *Oncotarget*. 7 (29), 45863–45875. doi:10.18632/oncotarget.10244.

Emert, B.L., Cote, C.J., Torre, E.A., Dardani, I.P., Jiang, C.L., Jain, N., Shaffer, S.M. & Raj, A. (2021) Variability within rare cell states enables multiple paths toward drug resistance. *Nature Biotechnology*. 39 (7), 865–876. doi:10.1038/s41587-021-00837-3.

Emran, A.A., Marzese, D.M., Menon, D.R., Stark, M.S., Torrano, J., et al. (2018) Distinct histone modifications denote early stress-induced drug tolerance in cancer. *Oncotarget*. 9 (9), 8206–8222. doi:10.18632/oncotarget.23654.

Ewels, P.A., Peltzer, A., Fillinger, S., Patel, H., Alneberg, J., Wilm, A., Garcia, M.U., Di Tommaso, P. & Nahnsen, S. (2020) The nf-core framework for community-curated bioinformatics pipelines. *Nature Biotechnology*. 38 (3), 276–278. doi:10.1038/s41587-020-0439-x.

Fan, J., Salathia, N., Liu, R., Kaeser, G.E., Yung, Y.C., Herman, J.L., Kaper, F., Fan, J.-B., Zhang, K., Chun, J. & Kharchenko, P.V. (2016) Characterizing transcriptional heterogeneity through pathway and gene set overdispersion analysis. *Nature Methods*. 13 (3), 241–244. doi:10.1038/nmeth.3734.

Fares, J., Fares, M.Y., Khachfe, H.H., Salhab, H.A. & Fares, Y. (2020) Molecular principles of metastasis: a hallmark of cancer revisited. *Signal Transduction and Targeted Therapy*. 5 (1), 28. doi:10.1038/s41392-020-0134-x.

Faure, A.J., Schmiedel, J.M. & Lehner, B. (2017) Systematic Analysis of the Determinants of Gene Expression Noise in Embryonic Stem Cells. *Cell Systems*. 5 (5), 471-484.e4. doi:10.1016/j.cels.2017.10.003.

Fennell, K.A., Vassiliadis, D., Lam, E.Y.N., Martelotto, L.G., Balic, J.J., Hollizeck, S., Weber, T.S., Semple, T., Wang, Q., Miles, D.C., MacPherson, L., Chan, Y.-C., Guirguis, A.A., Kats, L.M., Wong, E.S., Dawson, S.-J., Naik, S.H. & Dawson, M.A. (2022) Non-genetic determinants of malignant clonal fitness at single-cell resolution. *Nature*. 601 (7891), 125–131. doi:10.1038/s41586-021-04206-7.

Ferrara, N. (2005) VEGF as a Therapeutic Target in Cancer. *Oncology*. 69 (Suppl. 3), 11–16. doi:10.1159/000088479.

Flavahan, W.A., Drier, Y., Liau, B.B., Gillespie, S.M., Venteicher, A.S., Stemmer-Rachamimov, A.O., Suvà, M.L. & Bernstein, B.E. (2016) Insulator dysfunction and oncogene activation in IDH mutant gliomas. *Nature*. 529 (7584), 110–114. doi:10.1038/nature16490.

Flavahan, W.A., Gaskell, E. & Bernstein, B.E. (2017) Epigenetic plasticity and the hallmarks of cancer. *Science (New York, N.Y.)*. 357 (6348), eaal2380. doi:10.1126/science.aal2380.

Gameiro, P.A. & Struhl, K. (2018) Nutrient Deprivation Elicits a Transcriptional and Translational Inflammatory Response Coupled to Decreased Protein Synthesis. *Cell Reports*. 24 (6), 1415–1424. doi:10.1016/j.celrep.2018.07.021.

Garcia, M., Juhos, S., Larsson, M., Olason, P.I., Martin, M., Eisfeldt, J., DiLorenzo, S., Sandgren, J., Díaz De Ståhl, T., Ewels, P., Wirta, V., Nistér, M., Käller, M. & Nystedt, B. (2020) Sarek: A portable workflow for whole-genome sequencing analysis of germline and somatic variants. *F1000Research*. 9, 63. doi:10.12688/f1000research.16665.2.

Gerlinger, M., Horswell, S., Larkin, J., Rowan, A.J., Salm, M.P., et al. (2014) Genomic architecture and evolution of clear cell renal cell carcinomas defined by multiregion sequencing. *Nature Genetics*. 46 (3), 225–233. doi:10.1038/ng.2891.

Gerlinger, M., Rowan, A.J., Horswell, S., Math, M., Larkin, J., et al. (2012) Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. *The New England Journal of Medicine*. 366 (10), 883–892. doi:10.1056/NEJMoa1113205.

Gonzalez-Perez, A., Jene-Sanz, A. & Lopez-Bigas, N. (2013) The mutational landscape of chromatin regulatory factors across 4,623 tumor samples. *Genome Biology*. 14 (9), r106. doi:10.1186/gb-2013-14-9-r106.

Grün, D. (2020) Revealing dynamics of gene expression variability in cell state space. *Nature Methods*. 17 (1), 45–49. doi:10.1038/s41592-019-0632-3.

Gui, P. & Bivona, T.G. (2022) Evolution of metastasis: new tools and insights. *Trends in Cancer*. 8 (2), 98–109. doi:10.1016/j.trecan.2021.11.002.

Gupta, P.B., Fillmore, C.M., Jiang, G., Shapira, S.D., Tao, K., Kuperwasser, C. & Lander, E.S. (2011) Stochastic state transitions give rise to phenotypic equilibrium in populations of cancer cells. *Cell*. 146 (4), 633–644. doi:10.1016/j.cell.2011.07.026.

Han, M., Jia, L., Lv, W., Wang, L. & Cui, W. (2019) Epigenetic Enzyme Mutations: Role in Tumorigenesis and Molecular Inhibitors. *Frontiers in Oncology*. 9, 194. doi:10.3389/fonc.2019.00194.

Hanahan, D. (2022) Hallmarks of Cancer: New Dimensions. *Cancer Discovery*. 12 (1), 31–46. doi:10.1158/2159-8290.CD-21-1059.

Henser-Brownhill, T., Monserrat, J. & Scaffidi, P. (2017) Generation of an arrayed CRISPR-Cas9 library targeting epigenetic regulators: from high-content screens to in vivo assays. *Epigenetics*. 12 (12), 1065–1075. doi:10.1080/15592294.2017.1395121.

Hinohara, K., Wu, H.-J., Sébastien Vigneau, null, McDonald, T.O., Igarashi, K.J., et al. (2019) KDM5 Histone Demethylase Activity Links Cellular Transcriptomic Heterogeneity to Therapeutic Resistance. *Cancer Cell*. 35 (2), 330–332. doi:10.1016/j.ccell.2019.01.012.

Holliday, R. (1994) Epigenetics: an overview. *Developmental Genetics*. 15 (6), 453–457. doi:10.1002/dvg.1020150602.

Hong, S.P., Chan, T.E., Lombardo, Y., Corleone, G., Rotmensz, N., Bravaccini, S., Rocca, A., Pruneri, G., McEwen, K.R., Coombes, R.C., Barozzi, I. & Magnani, L. (2019) Single-cell transcriptomics reveals multi-step adaptations to endocrine therapy. *Nature Communications*. 10 (1), 3840. doi:10.1038/s41467-019-11721-9.

Hu, Z., Li, Z., Ma, Z. & Curtis, C. (2020) Multi-cancer analysis of clonality and the timing of systemic spread in paired primary tumors and metastases. *Nature Genetics*. 52 (7), 701–708. doi:10.1038/s41588-020-0628-z.

Huang, J., Gou, H., Yao, J., Yi, K., Jin, Z., Matsuoka, M. & Zhao, T. (2021) The noncanonical role of EZH2 in cancer. *Cancer Science*. 112 (4), 1376–1382. doi:10.1111/cas.14840.

Hugo, W., Shi, H., Sun, L., Piva, M., Song, C., Kong, X., Moriceau, G., Hong, A., Dahlman, K.B., Johnson, D.B., Sosman, J.A., Ribas, A. & Lo, R.S. (2015) Non-genomic and Immune Evolution of Melanoma Acquiring MAPKi Resistance. *Cell*. 162 (6), 1271–1285. doi:10.1016/j.cell.2015.07.061.

Iyer, N.G., Xian, J., Chin, S.-F., Bannister, A.J., Daigo, Y., Aparicio, S., Kouzarides, T. & Caldas, C. (2007) p300 is required for orderly G1/S transition in human cancer cells. *Oncogene*. 26 (1), 21–29. doi:10.1038/sj.onc.1209771.

Jamal-Hanjani, M., Wilson, G.A., McGranahan, N., Birkbak, N.J., Watkins, T.B.K., et al. (2017) Tracking the Evolution of Non–Small-Cell Lung Cancer. *New England Journal of Medicine*. 376 (22), 2109–2121. doi:10.1056/NEJMoa1616288.

Johnpulle, R.A.N., Johnson, D.B. & Sosman, J.A. (2016) Molecular Targeted Therapy Approaches for BRAF Wild-Type Melanoma. *Current Oncology Reports*. 18 (1), 6. doi:10.1007/s11912-015-0485-6.

Johnson, K.C., Anderson, K.J., Courtois, E.T., Gujar, A.D., Barthel, F.P., et al. (2021) Single-cell multimodal glioma analyses identify epigenetic regulators of cellular plasticity and environmental stress response. *Nature Genetics*. 53 (10), 1456–1468. doi:10.1038/s41588-021-00926-8.

Junttila, M.R. & de Sauvage, F.J. (2013) Influence of tumour micro-environment heterogeneity on therapeutic response. *Nature*. 501 (7467), 346–354. doi:10.1038/nature12626.

Kamphorst, J.J., Nofal, M., Commisso, C., Hackett, S.R., Lu, W., Grabocka, E., Vander Heiden, M.G., Miller, G., Drebin, J.A., Bar-Sagi, D., Thompson, C.B. & Rabinowitz, J.D. (2015) Human Pancreatic Cancer Tumors Are Nutrient Poor and Tumor Cells Actively Scavenge Extracellular Protein. *Cancer Research*. 75 (3), 544–553. doi:10.1158/0008-5472.CAN-14-2211.

Kantidakis, T., Saponaro, M., Mitter, R., Horswell, S., Kranz, A., Boeing, S., Aygün, O., Kelly, G.P., Matthews, N., Stewart, A., Stewart, A.F. & Svejstrup, J.Q. (2016) Mutation of cancer driver MLL2 results in transcription stress and genome instability. *Genes & Development*. 30 (4), 408–420. doi:10.1101/gad.275453.115.

Kanu, N., Grönroos, E., Martinez, P., Burrell, R.A., Yi Goh, X., et al. (2015) SETD2 loss-of-function promotes renal cancer branched evolution through replication stress and impaired DNA repair. *Oncogene*. 34 (46), 5699–5708. doi:10.1038/onc.2015.24.

Kapur, P., Peña-Llopis, S., Christie, A., Zhrebker, L., Pavía-Jiménez, A., Rathmell, W.K., Xie, X.-J. & Brugarolas, J. (2013) Effects on survival of BAP1 and PBRM1 mutations in sporadic clear-cell renal-cell carcinoma: a retrospective analysis with independent validation. *The Lancet. Oncology*. 14 (2), 159–167. doi:10.1016/S1470-2045(12)70584-3.

Karolchik, D. (2004) The UCSC Table Browser data retrieval tool. *Nucleic Acids Research*. 32 (90001), 493D – 496. doi:10.1093/nar/gkh103.

Karras, P., Bordeu, I., Pozniak, J., Nowosad, A., Pazzi, C., et al. (2022) A cellular hierarchy in melanoma uncouples growth and metastasis. *Nature*. 610 (7930), 190–198. doi:10.1038/s41586-022-05242-7.

Kaur, A., Cuenca, L., Kiani, K., Busch, G.T., Fingerman, D., Dunagin, M.C., Li, J., Dardani, I., Sanford, E.M., Pemberton, J., Goyal, Y., Weeraratna, A.T., Herlyn, M. & Raj, A. (2022) *Metastatic potential in clonal melanoma cells is driven by a rare, early-invading subpopulation*. doi:10.1101/2022.04.17.488591.

Kazanets, A., Shorstova, T., Hilmi, K., Marques, M. & Witcher, M. (2016) Epigenetic silencing of tumor suppressor genes: Paradigms, puzzles, and potential. *Biochimica Et Biophysica Acta*. 1865 (2), 275–288. doi:10.1016/j.bbcan.2016.04.001.

Kim, K.-P., Choi, J., Yoon, J., Bruder, J.M., Shin, B., Kim, J., Arauzo-Bravo, M.J., Han, D., Wu, G., Han, D.W., Kim, J., Cramer, P. & Schöler, H.R. (2021) Permissive epigenomes endow reprogramming competence to transcriptional regulators. *Nature Chemical Biology*. 17 (1), 47–56. doi:10.1038/s41589-020-0618-6.

Komurov, K., Tseng, J.-T., Muller, M., Seviour, E.G., Moss, T.J., Yang, L., Nagrath, D. & Ram, P.T. (2012) The glucose-deprivation network counteracts lapatinib-induced toxicity in resistant ErbB2-positive breast cancer cells. *Molecular Systems Biology*. 8, 596. doi:10.1038/msb.2012.25.

Kondo, H., Ratcliffe, C.D.H., Hooper, S., Ellis, J., MacRae, J.I., Hennequart, M., Dunsby, C.W., Anderson, K.I. & Sahai, E. (2021) Single-cell resolved imaging reveals intra-tumor heterogeneity in glycolysis, transitions between metabolic states, and their regulatory mechanisms. *Cell Reports*. 34 (7), 108750. doi:10.1016/j.celrep.2021.108750.

Kurdistani, S.K. (2007) Histone modifications as markers of cancer prognosis: a cellular view. *British Journal of Cancer*. 97 (1), 1–5. doi:10.1038/sj.bjc.6603844.

Landau, D.A., Carter, S.L., Stojanov, P., McKenna, A., Stevenson, K., et al. (2013) Evolution and impact of subclonal mutations in chronic lymphocytic leukemia. *Cell*. 152 (4), 714–726. doi:10.1016/j.cell.2013.01.019.

Larsson, A.J.M., Johnsson, P., Hagemann-Jensen, M., Hartmanis, L., Faridani, O.R., Reinius, B., Segerstolpe, Å., Rivera, C.M., Ren, B. & Sandberg, R. (2019) Genomic encoding of transcriptional burst kinetics. *Nature*. 565 (7738), 251–254. doi:10.1038/s41586-018-0836-1.

Lawrence, M.S., Stojanov, P., Mermel, C.H., Robinson, J.T., Garraway, L.A., Golub, T.R., Meyerson, M., Gabriel, S.B., Lander, E.S. & Getz, G. (2014) Discovery and saturation analysis of cancer genes across 21 tumour types. *Nature*. 505 (7484), 495–501. doi:10.1038/nature12912.

Lee, L.J., Papadopoli, D., Jewer, M., Del Rincon, S., Topisirovic, I., Lawrence, M.G. & Postovit, L.-M. (2021) Cancer Plasticity: The Role of mRNA Translation. *Trends in Cancer*. 7 (2), 134–145. doi:10.1016/j.trecan.2020.09.005.

Levine, J.H., Lin, Y. & Elowitz, M.B. (2013) Functional roles of pulsing in genetic circuits. *Science (New York, N.Y.)*. 342 (6163), 1193–1200. doi:10.1126/science.1239999.

Lewis, P.W., Müller, M.M., Koletsky, M.S., Cordero, F., Lin, S., Banaszynski, L.A., Garcia, B.A., Muir, T.W., Becher, O.J. & Allis, C.D. (2013) Inhibition of PRC2 activity by a gain-of-function H3 mutation found in pediatric glioblastoma. *Science (New York, N.Y.)*. 340 (6134), 857–861. doi:10.1126/science.1232245.

Lewis, S.M., Asselin-Labat, M.-L., Nguyen, Q., Berthelet, J., Tan, X., Wimmer, V.C., Merino, D., Rogers, K.L. & Naik, S.H. (2021) Spatial omics and multiplexed imaging to explore cancer biology. *Nature Methods*. 18 (9), 997–1012. doi:10.1038/s41592-021-01203-6.

Li, B. & Dewey, C.N. (2011) RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics*. 12 (1), 323. doi:10.1186/1471-2105-12-323.

Li, Z., Seehawer, M. & Polyak, K. (2022) Untangling the web of intratumour heterogeneity. *Nature Cell Biology*. 24 (8), 1192–1201. doi:10.1038/s41556-022-00969-x.

Liakos, A., Konstantopoulos, D., Lavigne, M.D. & Fousteri, M. (2020) Continuous transcription initiation guarantees robust repair of all transcribed genes and regulatory regions. *Nature Communications*. 11 (1), 916. doi:10.1038/s41467-020-14566-9.

Liang, Q., Kong, J., Stalker, J. & Bradley, A. (2009) Chromosomal mobilization and reintegration of Sleeping Beauty and PiggyBac transposons. *genesis*. 47 (6), 404–408. doi:10.1002/dvg.20508.

Liau, B.B., Sievers, C., Donohue, L.K., Gillespie, S.M., Flavahan, W.A., et al. (2017) Adaptive Chromatin Remodeling Drives Glioblastoma Stem Cell Plasticity and Drug Tolerance. *Cell Stem Cell*. 20 (2), 233-246.e7. doi:10.1016/j.stem.2016.11.003.

Limberger, T., Schlederer, M., Trachtová, K., Garces de Los Fayos Alonso, I., Yang, J., et al. (2022) KMT2C methyltransferase domain regulated INK4A expression suppresses prostate cancer metastasis. *Molecular Cancer*. 21 (1), 89. doi:10.1186/s12943-022-01542-8.

Lipinski, K.A., Barber, L.J., Davies, M.N., Ashenden, M., Sottoriva, A. & Gerlinger, M. (2016) Cancer Evolution and the Limits of Predictability in Precision Cancer Medicine. *Trends in Cancer*. 2 (1), 49–63. doi:10.1016/j.trecan.2015.11.003.

López-Maury, L., Marguerat, S. & Bähler, J. (2008) Tuning gene expression to changing environments: from rapid responses to evolutionary adaptation. *Nature Reviews. Genetics*. 9 (8), 583–593. doi:10.1038/nrg2398.

Loukas, I., Simeoni, F., Milan, M., Inglese, P., Patel, H., Goldstone, R., East, P., Strohbuecker, S., Mitter, R., Talsania, B., Tang, W., Ratcliffe, C.D.H., Sahai, E., Shahrezaei, V. & Scaffidi, P. (2023) Selective advantage of epigenetically disrupted cancer cells via phenotypic inertia. *Cancer Cell*. 41 (1), 70-87.e14. doi:10.1016/j.ccell.2022.10.002.

Love, M.I., Huber, W. & Anders, S. (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology*. 15 (12), 550. doi:10.1186/s13059-014-0550-8.

Makohon-Moore, A.P., Zhang, M., Reiter, J.G., Bozic, I., Allen, B., et al. (2017) Limited heterogeneity of known driver gene mutations among the metastases of individual patients with pancreatic cancer. *Nature Genetics*. 49 (3), 358–366. doi:10.1038/ng.3764.

Marine, J.-C., Dawson, S.-J. & Dawson, M.A. (2020) Non-genetic mechanisms of therapeutic resistance in cancer. *Nature Reviews Cancer*. 20 (12), 743–756. doi:10.1038/s41568-020-00302-4.

Marsolier, J., Prompsy, P., Durand, A., Lyne, A.-M., Landragin, C., et al. (2022) H3K27me3 conditions chemotolerance in triple-negative breast cancer. *Nature Genetics*. 54 (4), 459–468. doi:10.1038/s41588-022-01047-6.

Martin, M. (2011) Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal*. 17 (1), 10. doi:10.14806/ej.17.1.200.

Martincorena, I., Fowler, J.C., Wabik, A., Lawson, A.R.J., Abascal, F., Hall, M.W.J., Cagan, A., Murai, K., Mahbubani, K., Stratton, M.R., Fitzgerald, R.C., Handford, P.A., Campbell, P.J., Saeb-Parsy, K. & Jones, P.H. (2018) Somatic mutant clones colonize the human esophagus with age. *Science*. 362 (6417), 911–917. doi:10.1126/science.aau3879.

Martincorena, I., Raine, K.M., Gerstung, M., Dawson, K.J., Haase, K., Van Loo, P., Davies, H., Stratton, M.R. & Campbell, P.J. (2017) Universal Patterns of Selection in Cancer and Somatic Tissues. *Cell*. 171 (5), 1029-1041.e21. doi:10.1016/j.cell.2017.09.042.

Martincorena, I., Roshan, A., Gerstung, M., Ellis, P., Van Loo, P., McLaren, S., Wedge, D.C., Fullam, A., Alexandrov, L.B., Tubio, J.M., Stebbings, L., Menzies, A., Widaa, S., Stratton, M.R., Jones, P.H. & Campbell, P.J. (2015) High burden and pervasive positive selection of somatic mutations in normal human skin. *Science*. 348 (6237), 880–886. doi:10.1126/science.aaa6806.

Masel, J. & Siegal, M.L. (2009) Robustness: mechanisms and consequences. *Trends in genetics: TIG*. 25 (9), 395–403. doi:10.1016/j.tig.2009.07.005.

Mazor, T., Pankov, A., Johnson, B.E., Hong, C., Hamilton, E.G., et al. (2015) DNA Methylation and Somatic Mutations Converge on the Cell Cycle and Define Similar Evolutionary Histories in Brain Tumors. *Cancer Cell*. 28 (3), 307–317. doi:10.1016/j.ccell.2015.07.012.

Mazor, T., Pankov, A., Song, J.S. & Costello, J.F. (2016) Intratumoral Heterogeneity of the Epigenome. *Cancer Cell*. 29 (4), 440–451. doi:10.1016/j.ccell.2016.03.009.

McDonald, O.G., Li, X., Saunders, T., Tryggvadottir, R., Mentch, S.J., Warmoes, M.O., Word, A.E., Carrer, A., Salz, T.H., Natsume, S., Stauffer, K.M., Makohon-Moore, A., Zhong, Y., Wu, H., Wellen, K.E., Locasale, J.W., Iacobuzio-Donahue, C.A. & Feinberg, A.P. (2017) Epigenomic reprogramming during pancreatic cancer progression links anabolic glucose metabolism to distant metastasis. *Nature Genetics*. 49 (3), 367–376. doi:10.1038/ng.3753.

McGranahan, N. & Swanton, C. (2017) Clonal Heterogeneity and Tumor Evolution: Past, Present, and the Future. *Cell*. 168 (4), 613–628. doi:10.1016/j.cell.2017.01.018.

Mills, J.C., Stanger, B.Z. & Sander, M. (2019) Nomenclature for cellular plasticity: are the terms as plastic as the cells themselves? *The EMBO journal*. 38 (19), e103148. doi:10.15252/embj.2019103148.

Monserrat, J., Morales Torres, C., Richardson, L., Wilson, T.S., Patel, H., Domart, M.-C., Horswell, S., Song, O.-R., Jiang, M., Crawford, M., Bui, M., Dalal, Y. & Scaffidi, P. (2021) Disruption of the MSL complex inhibits tumour maintenance by exacerbating chromosomal instability. *Nature Cell Biology*. 23 (4), 401–412. doi:10.1038/s41556-021-00657-2.

Morris, L.G.T., Riaz, N., Desrichard, A., Şenbabaoğlu, Y., Hakimi, A.A., Makarov, V., Reis-Filho, J.S. & Chan, T.A. (2016) Pan-cancer analysis of intratumor heterogeneity as a prognostic determinant of survival. *Oncotarget*. 7 (9), 10051–10063. doi:10.18632/oncotarget.7067.

Mortimer, T., Wainwright, E.N., Patel, H., Siow, B.M., Jaunmuktane, Z., Brandner, S. & Scaffidi, P. (2019) Redistribution of EZH2 promotes malignant phenotypes by rewiring developmental programmes. *EMBO reports*. 20 (10), e48155. doi:10.15252/embr.201948155.

de Nadal, E., Ammerer, G. & Posas, F. (2011) Controlling gene expression in response to stress. *Nature Reviews. Genetics*. 12 (12), 833–845. doi:10.1038/nrg3055.

Nam, A.S., Chaligne, R. & Landau, D.A. (2021) Integrating genetic and non-genetic determinants of cancer evolution by single-cell multi-omics. *Nature Reviews. Genetics*. 22 (1), 3–18. doi:10.1038/s41576-020-0265-5.

Newman, J.R.S., Ghaemmaghami, S., Ihmels, J., Breslow, D.K., Noble, M., DeRisi, J.L. & Weissman, J.S. (2006) Single-cell proteomic analysis of S. cerevisiae reveals the architecture of biological noise. *Nature*. 441 (7095), 840–846. doi:10.1038/nature04785.

Nguyen, A., Yoshida, M., Goodarzi, H. & Tavazoie, S.F. (2016) Highly variable cancer subpopulations that exhibit enhanced transcriptome variability and metastatic fitness. *Nature Communications*. 7, 11246. doi:10.1038/ncomms11246.

Nieto, M.A., Huang, R.Y.-J., Jackson, R.A. & Thiery, J.P. (2016) EMT: 2016. *Cell*. 166 (1), 21–45. doi:10.1016/j.cell.2016.06.028.

Nowak, K., Seisenbacher, G., Hafen, E. & Stocker, H. (2013) Nutrient restriction enhances the proliferative potential of cells lacking the tumor suppressor PTEN in mitotic tissues. *eLife*. 2, e00380. doi:10.7554/eLife.00380.

Nowell, P.C. (1976) The clonal evolution of tumor cell populations. *Science (New York, N.Y.)*. 194 (4260), 23–28. doi:10.1126/science.959840.

Oakes, C.C., Claus, R., Gu, L., Assenov, Y., Hüllein, J., et al. (2014) Evolution of DNA methylation is linked to genetic aberrations in chronic lymphocytic leukemia. *Cancer Discovery*. 4 (3), 348–361. doi:10.1158/2159-8290.CD-13-0349.

O'Carroll, D., Scherthan, H., Peters, A.H., Opravil, S., Haynes, A.R., Laible, G., Rea, S., Schmid, M., Lebersorger, A., Jerratsch, M., Sattler, L., Mattei, M.G., Denny, P., Brown, S.D., Schweizer, D. & Jenuwein, T. (2000) Isolation and characterization of Suv39h2, a second histone H3 methyltransferase gene that displays testis-specific expression. *Molecular and Cellular Biology*. 20 (24), 9423–9433. doi:10.1128/MCB.20.24.9423-9433.2000.

Ochiai, H., Hayashi, T., Umeda, M., Yoshimura, M., Harada, A., Shimizu, Y., Nakano, K., Saitoh, N., Liu, Z., Yamamoto, T., Okamura, T., Ohkawa, Y., Kimura, H. & Nikaido, I. (2020) Genome-wide kinetic properties of transcriptional bursting in mouse embryonic stem cells. *Science Advances*. 6 (25), eaaz6699. doi:10.1126/sciadv.aaz6699.

Ortmann, C.A., Kent, D.G., Nangalia, J., Silber, Y., Wedge, D.C., et al. (2015) Effect of Mutation Order on Myeloproliferative Neoplasms. *New England Journal of Medicine*. 372 (7), 601–612. doi:10.1056/NEJMoa1412098.

Pakos-Zebrucka, K., Koryga, I., Mnich, K., Ljujic, M., Samali, A. & Gorman, A.M. (2016) The integrated stress response. *EMBO reports*. 17 (10), 1374–1395. doi:10.15252/embr.201642195.

Pan, M., Reid, M.A., Lowman, X.H., Kulkarni, R.P., Tran, T.Q., et al. (2016) Regional glutamine deficiency in tumours promotes dedifferentiation through inhibition of histone demethylation. *Nature Cell Biology*. 18 (10), 1090–1101. doi:10.1038/ncb3410.

Parker, H., Rose-Zerilli, M.J.J., Larrayoz, M., Clifford, R., Edelmann, J., et al. (2016) Genomic disruption of the histone methyltransferase SETD2 in chronic lymphocytic leukaemia. *Leukemia*. 30 (11), 2179–2186. doi:10.1038/leu.2016.134.

Pastore, A., Gaiti, F., Lu, S.X., Brand, R.M., Kulm, S., et al. (2019) Corrupted coordination of epigenetic modifications leads to diverging chromatin states and transcriptional heterogeneity in CLL. *Nature Communications*. 10 (1), 1874. doi:10.1038/s41467-019-09645-5.

Pflaum, J., Schlosser, S. & Müller, M. (2014) p53 Family and Cellular Stress Responses in Cancer. *Frontiers in Oncology*. 4, 285. doi:10.3389/fonc.2014.00285.

Piunti, A. & Shilatifard, A. (2021) The roles of Polycomb repressive complexes in mammalian development and cancer. *Nature Reviews Molecular Cell Biology*. 22 (5), 326–345. doi:10.1038/s41580-021-00341-1.

Priestley, P., Baber, J., Lolkema, M.P., Steeghs, N., de Bruijn, E., et al. (2019) Pan-cancer whole-genome analyses of metastatic solid tumours. *Nature*. 575 (7781), 210–216. doi:10.1038/s41586-019-1689-y.

Puig, I., Tenbaum, S.P., Chicote, I., Arqués, O., Martínez-Quintanilla, J., et al. (2018) TET2 controls chemoresistant slow-cycling cancer cell survival and tumor recurrence. *The Journal of Clinical Investigation*. 128 (9), 3887–3905. doi:10.1172/JCI96393.

Qing, G., Li, B., Vu, A., Skuli, N., Walton, Z.E., Liu, X., Mayes, P.A., Wise, D.R., Thompson, C.B., Maris, J.M., Hogarty, M.D. & Simon, M.C. (2012) ATF4 regulates MYC-mediated neuroblastoma cell death upon glutamine deprivation. *Cancer Cell*. 22 (5), 631–644. doi:10.1016/j.ccr.2012.09.021.

Quinn, J.J., Jones, M.G., Okimoto, R.A., Nanjo, S., Chan, M.M., Yosef, N., Bivona, T.G. & Weissman, J.S. (2021) Single-cell lineages reveal the rates, routes, and drivers of metastasis in cancer xenografts. *Science (New York, N.Y.)*. 371 (6532), eabc1944. doi:10.1126/science.abc1944.

Rambow, F., Rogiers, A., Marin-Bejar, O., Aibar, S., Femel, J., et al. (2018) Toward Minimal Residual Disease-Directed Therapy in Melanoma. *Cell*. 174 (4), 843-855.e19. doi:10.1016/j.cell.2018.06.025.

Ravindran Menon, D., Das, S., Krepler, C., Vultur, A., Rinner, B., Schauer, S., Kashofer, K., Wagner, K., Zhang, G., Bonyadi Rad, E., Haass, N.K., Soyer, H.P., Gabrielli, B., Somasundaram, R., Hoefler, G., Herlyn, M. & Schaider, H. (2015) A stress-induced early innate response causes multidrug tolerance in melanoma. *Oncogene*. 34 (34), 4448–4459. doi:10.1038/onc.2014.372.

Rawat, P., Boehning, M., Hummel, B., Aprile-Garcia, F., Pandit, A.S., Eisenhardt, N., Khavaran, A., Niskanen, E., Vos, S.M., Palvimo, J.J., Pichler, A., Cramer, P. & Sawarkar, R. (2021) Stress-induced nuclear condensation of NELF drives transcriptional downregulation. *Molecular Cell*. 81 (5), 1013-1026.e11. doi:10.1016/j.molcel.2021.01.016.

Rodriguez, J. & Larson, D.R. (2020) Transcription in Living Cells: Molecular Mechanisms of Bursting. *Annual Review of Biochemistry*. 89, 189–212. doi:10.1146/annurev-biochem-011520-105250.

Roe, J.-S., Hwang, C.-I., Somerville, T.D.D., Milazzo, J.P., Lee, E.J., et al. (2017) Enhancer Reprogramming Promotes Pancreatic Cancer Metastasis. *Cell*. 170 (5), 875-888.e20. doi:10.1016/j.cell.2017.07.007.

Rohani, N., Hao, L., Alexis, M.S., Joughin, B.A., Krismer, K., Moufarrej, M.N., Soltis, A.R., Lauffenburger, D.A., Yaffe, M.B., Burge, C.B., Bhatia, S.N. & Gertler, F.B. (2019) Acidification of Tumor at Stromal Boundaries Drives Transcriptome Alterations Associated with Aggressive Phenotypes. *Cancer Research*. 79 (8), 1952–1966. doi:10.1158/0008-5472.CAN-18-1604.

Saadatpour, A., Lai, S., Guo, G. & Yuan, G.-C. (2015) Single-Cell Analysis in Cancer Genomics. *Trends in genetics: TIG*. 31 (10), 576–586. doi:10.1016/j.tig.2015.07.003.

Sahai, E., Astsaturov, I., Cukierman, E., DeNardo, D.G., Egeblad, M., et al. (2020) A framework for advancing our understanding of cancer-associated fibroblasts. *Nature Reviews. Cancer*. 20 (3), 174–186. doi:10.1038/s41568-019-0238-1.

Salehi, S., Kabeer, F., Ceglia, N., Andronescu, M., Williams, M.J., et al. (2021) Clonal fitness inferred from time-series modelling of single-cell cancer genomes. *Nature*. 595 (7868), 585–590. doi:10.1038/s41586-021-03648-3.

Saunders, C.T., Wong, W.S.W., Swamy, S., Becq, J., Murray, L.J. & Cheetham, R.K. (2012) Strelka: accurate somatic small-variant calling from sequenced tumor–normal sample pairs. *Bioinformatics*. 28 (14), 1811–1817. doi:10.1093/bioinformatics/bts271.

Scaffidi, P. (2016) Histone H1 alterations in cancer. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms*. 1859 (3), 533–539. doi:10.1016/j.bbagrm.2015.09.008.

Shaffer, S.M., Dunagin, M.C., Torborg, S.R., Torre, E.A., Emert, B., Krepler, C., Beqiri, M., Sproesser, K., Brafford, P.A., Xiao, M., Eggan, E., Anastopoulos, I.N., Vargas-Garcia, C.A., Singh, A., Nathanson, K.L., Herlyn, M. & Raj, A. (2017) Rare cell variability and drug-induced reprogramming as a mode of cancer drug resistance. *Nature*. 546 (7658), 431–435. doi:10.1038/nature22794.

Shaffer, S.M., Emert, B.L., Reyes Hueros, R.A., Cote, C., Harmange, G., Schaff, D.L., Sizemore, A.E., Gupte, R., Torre, E., Singh, A., Bassett, D.S. & Raj, A. (2020) Memory Sequencing Reveals Heritable Single-Cell Gene Expression Programs Associated with Distinct Cellular Behaviors. *Cell*. 182 (4), 947-959.e17. doi:10.1016/j.cell.2020.07.003.

Sharma, S.V., Lee, D.Y., Li, B., Quinlan, M.P., Takahashi, F., Maheswaran, S., McDermott, U., Azizian, N., Zou, L., Fischbach, M.A., Wong, K.-K., Brandstetter, K., Wittner, B.,

Ramaswamy, S., Classon, M. & Settleman, J. (2010) A chromatin-mediated reversible drug-tolerant state in cancer cell subpopulations. *Cell*. 141 (1), 69–80. doi:10.1016/j.cell.2010.02.027.

Shen, H. & Laird, P.W. (2013) Interplay between the Cancer Genome and Epigenome. *Cell*. 153 (1), 38–55. doi:10.1016/j.cell.2013.03.008.

Shen, S., Vagner, S. & Robert, C. (2020) Persistent Cancer Cells: The Deadly Survivors. *Cell*. 183 (4), 860–874. doi:10.1016/j.cell.2020.10.027.

Simeoni, F., Loukas, I., Wilson, T.S. & Scaffidi, P. (2023) CRISPR-based large-scale modeling of loss-of-function mutations to investigate mechanisms of stress resistance in cancer. *STAR Protocols*. 4 (1), 102097. doi:10.1016/j.xpro.2023.102097.

Sottoriva, A., Kang, H., Ma, Z., Graham, T.A., Salomon, M.P., Zhao, J., Marjoram, P., Siegmund, K., Press, M.F., Shibata, D. & Curtis, C. (2015) A Big Bang model of human colorectal tumor growth. *Nature Genetics*. 47 (3), 209–216. doi:10.1038/ng.3214.

St Pierre, R. & Kadoch, C. (2017) Mammalian SWI/SNF complexes in cancer: emerging therapeutic opportunities. *Current Opinion in Genetics & Development*. 42, 56–67. doi:10.1016/j.gde.2017.02.004.

Sulli, G., Di Micco, R. & d'Adda di Fagagna, F. (2012) Crosstalk between chromatin state and DNA damage response in cellular senescence and cancer. *Nature Reviews. Cancer*. 12 (10), 709–720. doi:10.1038/nrc3344.

Sun, C., Wang, L., Huang, S., Heynen, G.J.J.E., Prahallad, A., et al. (2014) Reversible and adaptive resistance to BRAF(V600E) inhibition in melanoma. *Nature*. 508 (7494), 118–122. doi:10.1038/nature13121.

Suzuki, H., Aoki, K., Chiba, K., Sato, Y., Shiozawa, Y., et al. (2015) Mutational landscape and clonal architecture in grade II and III gliomas. *Nature Genetics*. 47 (5), 458–468. doi:10.1038/ng.3273.

Tajan, M., Hock, A.K., Blagih, J., Robertson, N.A., Labuschagne, C.F., Kruiswijk, F., Humpton, T.J., Adams, P.D. & Vousden, K.H. (2018) A Role for p53 in the Adaptation to Glutamine Starvation through the Expression of SLC1A3. *Cell Metabolism*. 28 (5), 721-736.e6. doi:10.1016/j.cmet.2018.07.005.

Tentler, J.J., Tan, A.C., Weekes, C.D., Jimeno, A., Leong, S., Pitts, T.M., Arcaroli, J.J., Messersmith, W.A. & Eckhardt, S.G. (2012) Patient-derived tumour xenografts as models for oncology drug development. *Nature Reviews. Clinical Oncology*. 9 (6), 338–350. doi:10.1038/nrclinonc.2012.61.

Torre, E.A., Arai, E., Bayatpour, S., Jiang, C.L., Beck, L.E., Emert, B.L., Shaffer, S.M., Mellis, I.A., Fane, M.E., Alicea, G.M., Budinich, K.A., Weeraratna, A.T., Shi, J. & Raj, A. (2021) Genetic screening for single-cell variability modulators driving therapy resistance. *Nature Genetics*. 53 (1), 76–85. doi:10.1038/s41588-020-00749-z.

Torrent, M., Chalancon, G., de Groot, N.S., Wuster, A. & Madan Babu, M. (2018) Cells alter their tRNA abundance to selectively regulate protein synthesis during stress conditions. *Science Signaling*. 11 (546), eaat6409. doi:10.1126/scisignal.aat6409.

Torres, C.M., Biran, A., Burney, M.J., Patel, H., Henser-Brownhill, T., Cohen, A.-H.S., Li, Y., Ben-Hamo, R., Nye, E., Spencer-Dene, B., Chakravarty, P., Efroni, S., Matthews, N., Misteli, T., Meshorer, E. & Scaffidi, P. (2016) The linker histone H1.0 generates epigenetic and functional intratumor heterogeneity. *Science (New York, N.Y.)*. 353 (6307), aaf1644. doi:10.1126/science.aaf1644.

Tufegdžić Vidaković, A., Mitter, R., Kelly, G.P., Neumann, M., Harreman, M., Rodríguez-Martínez, M., Herlihy, A., Weems, J.C., Boeing, S., Encheva, V., Gaul, L., Milligan, L., Tollervey, D., Conaway, R.C., Conaway, J.W., Snijders, A.P., Stewart, A. & Svejstrup, J.Q. (2020) Regulation of the RNAPII Pool Is Integral to the DNA Damage Response. *Cell*. 180 (6), 1245-1261.e21. doi:10.1016/j.cell.2020.02.009.

Tulstrup, M., Soerensen, M., Hansen, J.W., Gillberg, L., Needhamsen, M., Kaastrup, K., Helin, K., Christensen, K., Weischenfeldt, J. & Grønbæk, K. (2021) TET2 mutations are associated with hypermethylation at key regulatory enhancers in normal and malignant hematopoiesis. *Nature Communications*. 12 (1), 6061. doi:10.1038/s41467-021-26093-2.

Tunnacliffe, E. & Chubb, J.R. (2020) What Is a Transcriptional Burst? *Trends in Genetics*. 36 (4), 288–297. doi:10.1016/j.tig.2020.01.003.

Turajlic, S., Sottoriva, A., Graham, T. & Swanton, C. (2019) Resolving genetic heterogeneity in cancer. *Nature Reviews. Genetics*. 20 (7), 404–416. doi:10.1038/s41576-019-0114-6.

Umkehrer, C., Holstein, F., Formenti, L., Jude, J., Froussios, K., Neumann, T., Cronin, S.M., Haas, L., Lipp, J.J., Burkard, T.R., Fellner, M., Wiesner, T., Zuber, J. & Obenauf, A.C. (2021)

Isolating live cell clones from barcoded populations using CRISPRa-inducible reporters. *Nature Biotechnology*. 39 (2), 174–178. doi:10.1038/s41587-020-0614-0.

Van Der Auwera, G.A., Carneiro, M.O., Hartl, C., Poplin, R., del Angel, G., Levy-Moonshine, A., Jordan, T., Shakir, K., Roazen, D., Thibault, J., Banks, E., Garimella, K.V., Altshuler, D., Gabriel, S. & DePristo, M.A. (2013) From FastQ Data to High-Confidence Variant Calls: The Genome Analysis Toolkit Best Practices Pipeline. *Current Protocols in Bioinformatics*. 43 (1). doi:10.1002/0471250953.bi1110s43.

Vasimuddin, Md., Misra, S., Li, H. & Aluru, S. (2019) Efficient Architecture-Aware Acceleration of BWA-MEM for Multicore Systems. In: *2019 IEEE International Parallel and Distributed Processing Symposium (IPDPS)*. May 2019 Rio de Janeiro, Brazil, IEEE. pp. 314–324. doi:10.1109/IPDPS.2019.00041.

Vendramin, R., Litchfield, K. & Swanton, C. (2021) Cancer evolution: Darwin and beyond. *The EMBO journal*. 40 (18), e108389. doi:10.15252/embj.2021108389.

Waddington, C.H. (2012) The epigenotype. 1942. *International Journal of Epidemiology*. 41 (1), 10–13. doi:10.1093/ije/dyr184.

Wainwright, E.N. & Scaffidi, P. (2017) Epigenetics and Cancer Stem Cells: Unleashing, Hijacking, and Restricting Cellular Plasticity. *Trends in Cancer*. 3 (5), 372–386. doi:10.1016/j.trecan.2017.04.004.

Wang, T., Wei, J.J., Sabatini, D.M. & Lander, E.S. (2014) Genetic Screens in Human Cells Using the CRISPR-Cas9 System. *Science*. 343 (6166), 80–84. doi:10.1126/science.1246981.

Wei, R., Liu, S., Zhang, S., Min, L. & Zhu, S. (2020) Cellular and Extracellular Components in Tumor Microenvironment and Their Application in Early Diagnosis of Cancers. *Analytical Cellular Pathology (Amsterdam)*. 2020, 6283796. doi:10.1155/2020/6283796.

Wu Ct, null & Morris, J.R. (2001) Genes, genetics, and epigenetics: a correspondence. *Science (New York, N.Y.)*. 293 (5532), 1103–1105. doi:10.1126/science.293.5532.1103.

Wunderlich, V. (2002) JMM---past and present. Chromosomes and cancer: Theodor Boveri's predictions 100 years later. *Journal of Molecular Medicine (Berlin, Germany)*. 80 (9), 545–548. doi:10.1007/s00109-002-0374-y.

Yachida, S., Jones, S., Bozic, I., Antal, T., Leary, R., Fu, B., Kamiyama, M., Hruban, R.H., Eshleman, J.R., Nowak, M.A., Velculescu, V.E., Kinzler, K.W., Vogelstein, B. & Iacobuzio-

Donahue, C.A. (2010) Distant metastasis occurs late during the genetic evolution of pancreatic cancer. *Nature*. 467 (7319), 1114–1117. doi:10.1038/nature09515.

Yang, C., Tian, C., Hoffman, T.E., Jacobsen, N.K. & Spencer, S.L. (2021) Melanoma subpopulations that rapidly escape MAPK pathway inhibition incur DNA damage and rely on stress signalling. *Nature Communications*. 12 (1), 1747. doi:10.1038/s41467-021-21549-x.

Yang, D., Jones, M.G., Naranjo, S., Rideout, W.M., Min, K.H.J., et al. (2022) Lineage tracing reveals the phylodynamics, plasticity, and paths of tumor evolution. *Cell*. 185 (11), 1905-1923.e25. doi:10.1016/j.cell.2022.04.015.

Yang, L., Rau, R. & Goodell, M.A. (2015) DNMT3A in haematological malignancies. *Nature Reviews. Cancer*. 15 (3), 152–165. doi:10.1038/nrc3895.

Yoo, H.C., Yu, Y.C., Sung, Y. & Han, J.M. (2020) Glutamine reliance in cell metabolism. *Experimental & Molecular Medicine*. 52 (9), 1496–1516. doi:10.1038/s12276-020-00504-8.

Zahir, N., Sun, R., Gallahan, D., Gatenby, R.A. & Curtis, C. (2020) Characterizing the ecological and evolutionary dynamics of cancer. *Nature Genetics*. 52 (8), 759–767. doi:10.1038/s41588-020-0668-4.

Zhang, J., Fan, J., Venneti, S., Cross, J.R., Takagi, T., Bhinder, B., Djaballah, H., Kanai, M., Cheng, E.H., Judkins, A.R., Pawel, B., Baggs, J., Cherry, S., Rabinowitz, J.D. & Thompson, C.B. (2014) Asparagine Plays a Critical Role in Regulating Cellular Adaptation to Glutamine Depletion. *Molecular Cell*. 56 (2), 205–218. doi:10.1016/j.molcel.2014.08.018.

Zhang, J., Pavlova, N.N. & Thompson, C.B. (2017) Cancer cell metabolism: the essential role of the nonessential amino acid, glutamine. *The EMBO journal*. 36 (10), 1302–1315. doi:10.15252/embj.201696151.

Appendix 1

Contributions to published work

The data generated during this PhD thesis have contributed to the two following publications:

I. Loukas, I., Simeoni, F., Milan, M., Inglese, P., Patel, H., Goldstone, R., East, P., Strohbuecker, S., Mitter, R., Talsania, B., Tang, W., Ratcliffe, C.D.H., Sahai, E., Shahrezaei, V. & Scaffidi, P. (2023) Selective advantage of epigenetically disrupted cancer cells via phenotypic inertia. Cancer Cell. 41 (1), 70-87.e14. doi:10.1016/j.ccell.2022.10.002.

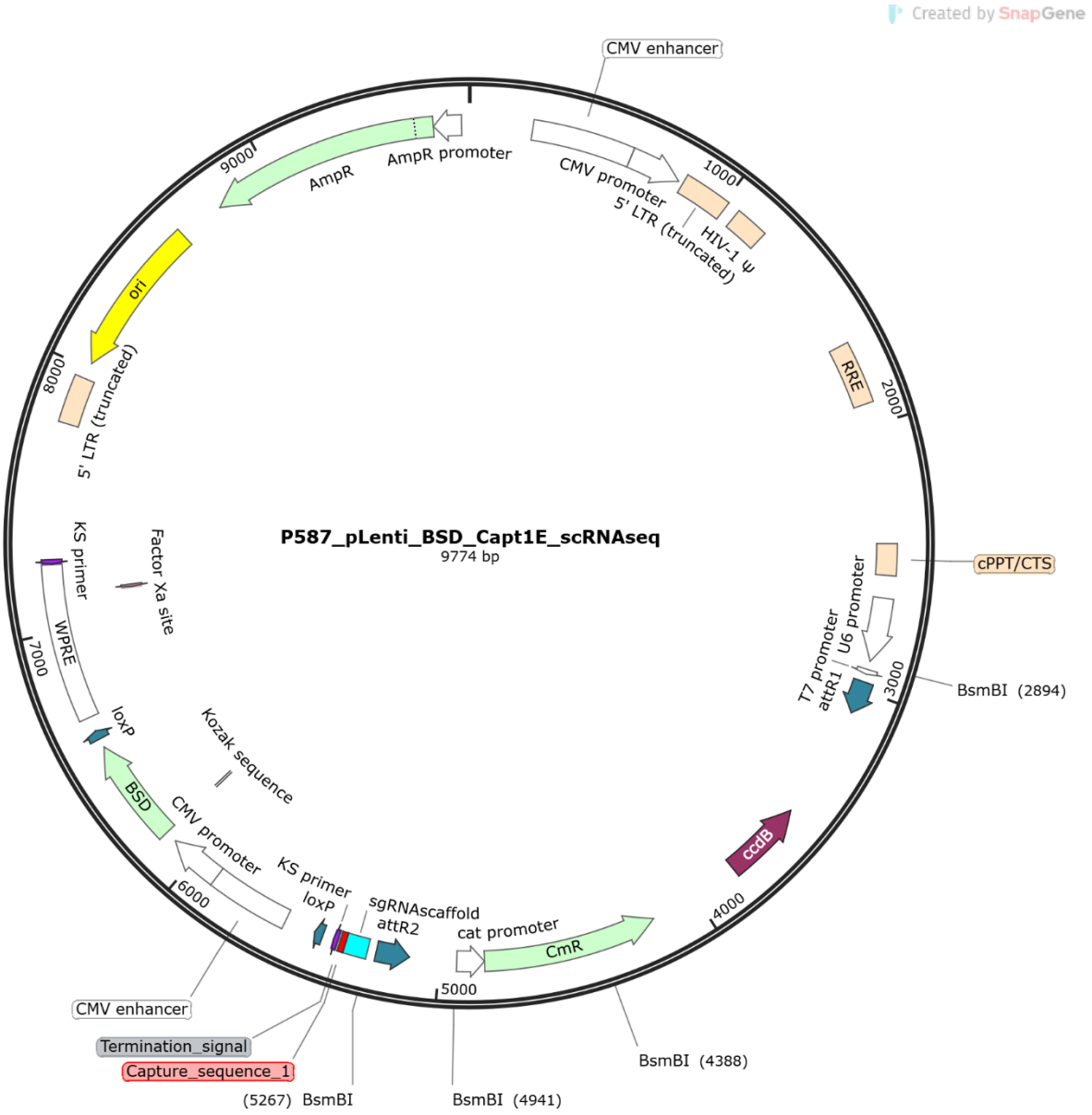
II. Simeoni, F., Loukas, I., Wilson, T.S. & Scaffidi, P. (2023) CRISPR-based large-scale modeling of loss-of-function mutations to investigate mechanisms of stress resistance in cancer. STAR Protocols. 4 (1), 102097. doi:10.1016/j.xpro.2023.102097.

Appendix 2

Maps of plasmids generated in this thesis or provided by other labs

P723_pTRIP-SFFV-mCherry-NLS (see Methods section 2.3.1)

P587_pLenti_BSD_Capt1E_scRNAseq (see Methods section 2.3.2)



pSBbi plasmid carrying the glucose FRET biosensor (from Sahai lab, see Methods section 2.14.1)