A Tale of Two Tumors: Unraveling the Complexities of Epigenetic Dysregulation in Cancer

Thèse N° 7228

Présentée le 23 août 2019

à la Faculté des sciences de la vie
Unité de la Prof. Oricchio
Programme doctoral en approches moléculaires du vivant

pour l'obtention du grade de Docteur ès Sciences

par

Maria Christine DONALDSON

Acceptée sur proposition du jury

Prof. J. Lingner, président du jury
Prof. E. Oricchio, directrice de thèse
Prof. N. Riggi, rapporteur
Dr L. Giorgetti, rapporteur
Prof. M. De Palma, rapporteur

2019

ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

“The greatest enemy of knowledge is not ignorance, it is the illusion of knowledge.”

Daniel Boorstin and Stephen Hawking

Acknowledgements

I would like to thank Elisa Oricchio for the opportunity to perform my doctoral studies under her supervision and for her mentorship over these last four years. I am also thanksful for the learning experience that comes with setting up a new laboratory, new techniques, and navigating a new institute. Most importantly, for showing me how to explore and tackle the difficult questions in science, and how to approach science from many different angles. Thank you for letting me work on the "risky" projects, and supporting them all the way to manuscript fruition.

I would like to thank Joachim Lingner, Michele De Palma, Nicolo Riggi, and Luca Giorgetti for participating on my Ph.D. defense committee.

I would like to thank Doug Hanahan for his guidance, mentorship, and support during my Ph.D. studies. I have valued our conversations and his help navigating every aspect of academia, from research, to career planning.

I would like to thank the members of the Oricchio lab for a nice working environment. Particularly for the helpful discussions, for laughing with me when experiments become disasters, and for finding time between experiments for coffee breaks and snacks.

I would like to thank the members of the Ciriello lab for their fantastic collaborations. In particular, Giovanni, Marco, and Sadegh for taking the time to help me with the many complex data analyses that came with my projects, and most importantly, for teaching me along the way.

I would like to thank Etienne Meylan and Joerg Huelsken for all of our fruitful and engaging conversations about my projects over the last four years. I most appreciate that you challenged me and asked me the difficult questions that encouraged me to think outside of the box.

I would like to thank all of our collaborators and the EPFL core facilities for your help in designing and executing our experiments. Your expertise and contributions were invaluable to our findings. Particularly, Kyle, Timo, and Suliana for your help with the STORM imaging experiments. Mor for your help with the metabolic imaging experiments. Jessica, Gianni, and Nathalie of the histology core facility. Loïc, Val, and Andre of the FACs core facility. Olivier, Romain, Thierry, and Arne of the BIOP Imaging facility. Bastien, Elisa and Lionel from the GECF sequencing facility.

I would like to thank Kathleen, Radek, Caterina, Isabelle, Simone, Selene, and Lucie for all the amazing adventures together. Between skiing, hiking, lake swims, concerts, dinners together and so much more, you made my time in Switzerland unforgettable.

I would like to thank my family and friends, from all over the world, for your consistent support and love. Particularly my parents and Edith - it really takes a village! The journey is much easier knowing that I have such amazing people cheering me on every step of the way.

To leave the best for last, Jake: thank you for being my sidekick, my cheerleader, my bad influence, my good influence, my favorite travel buddy, and love of my life. The adventure is infinitely better with you by my side.

Abstract

Epigenetics plays an important role in cancer development and progression. Cancer cells hijack the epigenome by modifying the histone protein units responsible for packaging DNA, or by modifying the DNA itself, resulting in changes to chromatin topology and transcriptional programming within the cell. In this thesis, I report our investigation to uncover the mechanisms of epigenetic regulation and how epigenetic control of transcriptional programming goes awry in cancer. We investigate these processes in two separate contexts.

In our first study, we investigate the mechanisms by which EZH2 oncogenic mutations alter structure and function of topologically associating domains in non-Hodgkin lymphoma.
The process of chromatin folding leads to a systematic arrangement of hierarchical structural and regulatory elements found within the nucleus. A topologically associating domain (TAD) is a regulatory unit of self interacting DNA, where the transcription of the genes within a TAD is usually dictated by the presence of active (H3K36me3) or inactive (H3K27me3) histone marks. In cancer, the somatic point mutation in $EZH2^{Y646X}$ leads to a genome wide increase in H3K27me3, associated with transcriptional repression. While this alteration leads to changes in the global epigenetic status of the cell, its impact on chromatin organization and TAD-related function remain unclear. By combining transcriptomics and epigenetics with analyses of chromatin structure, we demonstrate a functional interplay between TADs and the epigenetic and transcriptional program of the genes found within them. This balance is altered by $EZH2^{Y646X}$, leading to the synergistic silencing of entire domains directly targeting cell differentiation and tumor suppressive programs. A closer look reveals that the silencing of tumor suppressive TADs are coupled with structural modifications and changes in promoter interactions within $EZH2^{Y646X}$ target TAD6.139. Impressively, the TAD's transcriptional and epigenetic programs are restored by pharmacological inhibition of $EZH2^{Y646X}$. Our results indicate that $EZH2^{Y646X}$ alters the topology and function of chromatin domains to promote synergistic oncogenic programs.

In our second study, we dissect the epigenetic, transcriptomic, and metabolic signaling dependencies using a novel model of central nervous system primitive neural ectodermal tumors (CNS-PNETs).
Central nervous system (CNS) tumors are the leading cause of cancer-associated death in children. Primitive neural-ectodermal tumors (CNS-PNETs) are a particularly aggressive subtype of embryonal CNS-tumor, with a five-year overall survival rate in less than 50% of patients. Despite sharing a similar cell of origin of other CNS tumors, CNS-PNETs have a different anatomical location, unique genetic and epigenetic features, and significantly worse clinical outcome. In addition, a lack of *in vivo* models for studying CNS tumors challenges the opportunity to dissect the genetic variables that underlie the origin of these tumors.
We developed a novel CNS-PNET mouse model, called CNS-NPCs, using neural progenitor cells derived from human-iPS cells. Through histological, DNA methylation, and RNA sequencing analyses, we find that the CNS-NPC model recapitulates the the morphologic, epigenetic, and transcriptomic features of primary CNS-PNETs. In addition, through *in vivo* metabolic analyses of CNS-NPCs and the patient derived cell line, PFSK-1, we identified dysregulation in the neurotransmitter associated metabolites, N-acetyl aspartate (NAA) and γ-amino butyric acid (GABA). In addition, we identified metabolic features associated to CNS-PNET tumors that can be used as potential biomarkers for diagnosis, including glycine, taurine, and choline. Finally, using *in vivo* ^{1}H NMR and ^{13}C hyperpolarized magnetic resonance imaging, we identified two metabolic biomarkers, myo-inositol and lactate, that correlate to tumor growth, aggressiveness and blood brain barrier damage.

Keywords:
Cancer epigenetics, Cancer metabolism, EZH2 mutations, Chromatin organization, Lymphoma, CNS-PNET

Resume

L'épigénétique joue un rôle important dans le développement et la progression du cancer. Les cellules cancéreuses détournent l'épigénome en modifiant les unités protéiques de l'histone responsables de l'emballage de l'ADN ou en modifiant l'ADN lui-même, ce qui entraîne des modifications de la topologie de la chromatine et de la programmation transcriptionnelle au sein de la cellule. Dans cette thèse, nous rapportons notre enquête pour découvrir les mécanismes de la régulation épigénétique et comment le contrôle épigénétique de la programmation transcriptionnelle va à l'encontre du cancer. Nous étudions ces processus dans deux contextes distincts.

Dans notre première étude, nous avons étudié les mécanismes par lesquels les mutations oncogéniques EZH2 modifient la structure et la fonction des domaines topologiquement associés dans les lymphomes non hodgkiniens.
Le processus de repliement de la chromatine conduit à un arrangement systématique des éléments structurels et régulateurs hiérarchiques présents dans le noyau. Un domaine topologiquement associant (TAD) est une unité régulatrice de l'ADN auto-interagissant, où la transcription des gènes au sein d'un TAD est généralement dictée par la présence de marques histones actives (H3K36me3) ou inactives (H3K27me3). Dans le cancer, la mutation ponctuelle somatique chez $EZH2^{Y646X}$ entraîne une augmentation du génome de H3K27me3, associée à une répression de la transcription. Bien que cette altération génère une variabilité dans le statut épigénétique global de la cellule, son impact sur l'organisation de la chromatine et sur la fonction liée à la TAD reste flou. En combinant transcriptomique et épigénétique avec des analyses de la structure de la chromatine, nous démontrons une interaction fonctionnelle entre les TAD et le programme épigénétique et transcriptionnel des gènes qui s'y trouvent. $EZH2^{Y646X}$ modifie cet équilibre, ce qui conduit à la désactivation synergique de domaines entiers ciblant directement les programmes spécifiques à la différenciation cellulaire et à la suppression des tumeurs. Un examen plus attentif révèle que le silence des TAD suppresseurs de tumeurs est associé à des modifications structurelles et à des modifications des interactions des promoteurs au sein du TAD. De manière impressionnante, les programmes transcriptionnels et épigénétiques du TAD sont restaurés par inhibition pharmacologique de $EZH2^{Y646X}$. Nos résultats indiquent que $EZH2^{Y646X}$ modifie la topologie et la fonction des domaines de la chromatine afin de promouvoir des programmes oncogéniques synergiques.

Dans notre deuxième étude, nous avons disséqué les dépendances de signalisation épigénétiques, transcriptomiques et métaboliques en utilisant un nouveau modèle de tumeurs ectodermiques neurales primitives du système nerveux central (CNS-PNET).
Les tumeurs du système nerveux central (SNC) sont la principale cause de décès associé au cancer chez les enfants. Les tumeurs neuro-ectodermiques primitives (SNC-PNET) sont un sous-type particulièrement agressif de tumeur embryonnaire du SNC, avec un taux de survie global à cinq ans chez moins de 50% des patients. Bien qu'ils partagent une cellule d'origine similaire à celle d'autres tumeurs du SNC, les systèmes CNS-PNET ont une localisation anatomique différente, des caractéristiques génétiques uniques et un résultat clinique nettement pire. En outre, l'absence de modèles *in vivo* pour l'étude des tumeurs du SNC remet en cause la possibilité de disséquer les variables génétiques qui sous-tendent l'origine de ces tumeurs.
Nous avons développé un nouveau modèle de souris CNS-PNET, appelé CNS-NPC, utilisant des cellules progénitrices neurales dérivées de cellules iPS humaines. Par des analyses histologiques, de méthylation de l'ADN et de séquençage d'ARN, nous trouvons que le modèle CNS-NPC récapitule les caractéristiques morphologiques, épigénétiques et transcriptomiques des systèmes primaires CNS-PNET. De plus, par des analyses métaboliques *in vivo* de NPC-CNS et de la lignée cellulaire dérivée du patient, PFSK-1, nous avons identifié une dysrégulation des métabolites associés aux neurotransmetteurs, l'aspartate de N-acétyle (NAA) et l'acide amino-butyrique (GABA). De plus, nous avons identifié des caractéristiques métaboliques associées aux tumeurs CNS-PNET pouvant être utilisées comme biomarqueurs potentiels pour le diagnostic, notamment la glycine, la taurine et la choline. Enfin, en utilisant l'imagerie par résonance magnétique 1H RMN et 13C hyperpolarisée *in vivo*, nous avons

identifié deux biomarqueurs métaboliques, le myo-inositol et le lactate, qui sont corrélés à la croissance tumorale, à l'agressivité et aux dommages de la barrière hémato-encéphalique.

Mots clés:
épigénétique du cancer, métabolisme du cancer, mutations EZH2, organisation de la chromatine, lymphome, CNS-PNET

Abbreviations

ABC - Activated B-Cell
Acetyl co-A - Acetyl coenzyme-A
AEBP2 - AE Binding Protein 2
Akt - Protein Kinase B
APC - Adenomatous polyposis coli
AML - Acute Myeloid Leukemia
ATP - Adenosine Triphosphate
AR- Androgen Receptor
BBB - Blood Brain Barrier
BLI - Bioluminesence Imaging
CDKN2 - Cyclin Dependent Kinase Inhibitor-2
CGI - CpG Island
CNS - Central Nervous System
CTCF -CCCTC-binding factor
DLBCL - Diffuse Large B-Cell Lymphoma
DNA - Deoxyribonucleic Acid
DNMT - DNA Methyltransferase
EED -Embryonic Ectoderm Development
EPN - Ependymyoma
ER-α - Estrogen Receptor-α
ESC - Embryonic Stem Cell
EZH2 - Enhancer Zeste Homologue-2
F - Phenylalanine
FDG - ^{18}fluorodeoxyglucose
FOXO3 -Forkhead Box O3
GABA - γ-amino butyric acid
GC - Germinal Center
GCB - Germinal Center B-Cell
GFAP -Glial Fibrillary Acidic Protein
GFP - Green Fluoresence Protein
HAT - Histone Acetyltransferase
HDAC - Histone Deacetylase
HGG - High Grade Glioma
Hi-C - High Throughput Chromatin Conformation Capture
HL - Hodgkin Lymhpoma
HMT - Histone Methyltransferase
HMR - Hyperpolarized Magnetic Resonance
IDH - Isocitrate Dehydrogenase
IPS cell - Induced Pluripotent Stem cell
JAK/STAT - Janus Kinase/Signal Transducers and Activators of Transcription
JARID2 - Jumonji And AT-Rich Interaction Domain Containing-2
KLF4 -Kruppel Like Factor 4
LC/MS - Liquid Chromatography Mass Spectrometry
MDS - Myelodysplastic Syndrome
mTOR - Mammalian Target of Rapamycin
N - Asparagine
NAA - N-Acetyl Aspartate
NAAG - N-acetyl-aspartyl-glutamate
Nestin - Neural Development Type-VI Intermediate Filament Protein
NHL -Non-Hodgkin Lymphoma
NMR - Nuclear Magnetic Resonance
NPC -Nerual Progenitor Cell
NOS - Not Otherwise Specified
MM - Macromolecules
MRI - Magnetic Resonance Imaging
OCT4 - Octamer-Binding Transcription Factor 4
OLIG1/2 - Oligodendrocyte Transcription Factor 1/2
p53 - Tumor Protein-53
PAX6 - Paired box protein-6
PcG - Polycomb Group
PCL - Polycomb-Like
PDGFR - platelet derived growth factor receptor
PET - Positron Emission Tomography
PI3K - Phosphatidylinositol 3-Kinase
PNET - Primitive Neural Ectodermal Tumor
PRC - Polycomb Repressive Complex
PRE - Polycomb Responsive Element
PTEN - Phosphatase and Tensin Homologue
RNA - Ribonucleic Acid
RASSF1 - Ras-Associated Domain Family 1
RbAp46/48-Retinoblastoma protein associated protein
SESN1 - Sestrin1
SNAI2 - Snail Family Transcriptional Repressor 2
SOX2 - SRY Box 2
SUZ12 -Suppressor of Zeste-12
TAD - Topologically Associating Domain
t-ALL - T-cell Acute Lymphocytic Leukemia
TrxG - Trithorax Group
TWST1 - Twist Basic Helix-Loop-Helix Transcription Factor 1
UMI - Unique Molecular Identifier
Y - Tyrosine
YY1 - Yin Yang 1

Contents

Chapter 3: Dissecting epigenetic, transcriptomic, and metabolic signaling dependencies using a novel model of central nervous system primitive neural ectodermal tumors (CNS-PNETs)

Chapter 4. Materials and Methods

List of Figures

Chapter 1: Introduction

Chapter 2: EZH2 oncogenic mutations drive epigenetic, transcriptional, and structural changes within chromatin domains

Chapter 3: Dissecting epigenetic, transcriptomic, and metabolic signaling dependencies using a novel model of central nervous system primitive neural ectodermal tumors (CNS-PNETs)

List of Tables

Chapter 4: Methods
Dissecting epigenetic, transcriptomic, and metabolic signaling dependencies using a novel model of central nervous system primitive neural ectodermal tumors (CNS-PNETs)

Acknowledgement of Contribution to Thesis Content

Figure1: "The Hallmarks of Cancer"
Complete figure obtained from *Cell* - with permission

Figure 2: "The PRC2 Complex"
Complete illustration provided by Olaneli - with permission

Figure 3: "Molecular Classification of CNS-PNETs by DNA Methylation Profiling"
Complete figure obtained from *Cell* - with permission

Figure 4: "Schematic representation of the principal pathways involved in glucose catabolism"
Complete figure obtained from *Frontiers in Immunology* - with permission

Figure 5. An *in vivo* Model for Studying CNS Tumors
Figure 5F and 5H: Immunostain was conducted by the histology core facility, EPFL

Figure 6. Analysis of DNA Metyhlation and RNA Expression in Pediatric CNS Tumors
Figure 6A, 6B: Methylation analysis of CNS-PNET tumors was completed by M. Mina (Ciriello Lab, Unil)
Figure 6C, 6D: Expression analysis of CNS-PNET tumors was completed by S. Saghafinia (Ciriello Lab, Unil; Hanahan Lab, EPFL)
Supplemental Figure 6B, C: Methylation analysis of CNS-PNET tumors was completed by M. Mina (Ciriello Lab, Unil)

Figure 7. *Ex vivo* Analysis of Glucose Metabolism
Figure 7A: Hydrophilic interaction liquid chromatography coupled to tandem mass spectrometry (HILIC - MS/MS) was conducted by the Ivanisevic Lab (Unil)

Figure 9. *In vivo* Metabolic Characterization of CNS-NPC and PFSK-1 Tumors
Figure 9: All ^{1}H and ^{13}C spectroscopy data were acquired and analyzed by M. Mishkovsky (Laboratory for functional and Metabolic Imaging, Rolf Gruetter, EPFL)

Elisa Oricchio assited in the experimental design and data interpretation

The author (M.Donaldson) is otherwise responsible for the quality of content presented herein with regard to experimental design, data acquisition, analysis, interpretation and visualization

Chapter 1: Introduction

"Down to their innate molecular core, cancer cells are hyperactive, survival-endowed, scrappy, fecund, inventive copies of ourselves."
The Emperor of all Maladies – Siddahartha Mukherjee

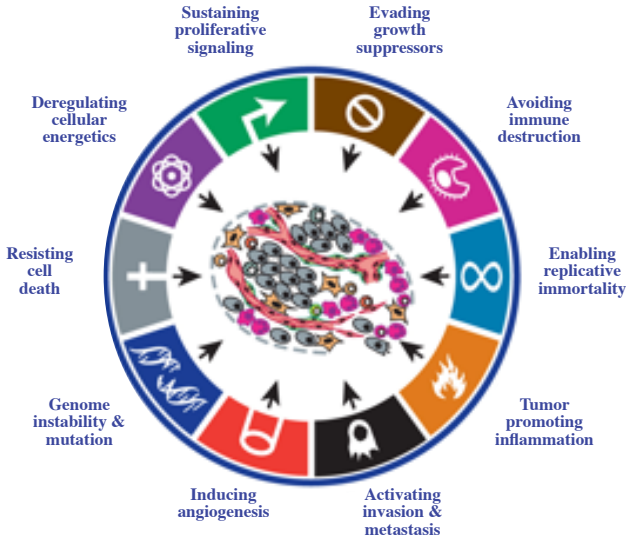
1.1 Cancer

Cancer is the second leading cause of death worldwide, with approximately 18 million new cases arising annually, and 9.8 million deaths recorded in 2018 alone (WHO). According to the 2018 WHO report, one in five men and one in six women will develop cancer during their lifetime. One in eight men and one in eleven women will die from the disease.

Cancer arises from a cell that acquires multiple alterations, where this accumulation over time eventually leads to the transformation from a normal cell to a cell that no longer retains control over it's own cellular processes, leading to rapid division and cell growth. These alterations can occur either by external factors, known as carcinogens, or by random errors that occur during DNA replication, and these alterations damage the cell's DNA beyond repair. Accumulation of DNA damage over time can result in a cell's transformation from normal to malignant.

Cancer is a complex disease with many types of alterations contributing to pathogenesis. Classically, cancer is generally defined by the cell or tissue of origin, although the complexity of alterations between and even within patients is highly heterogeneous, with each tumor exhibiting a unique assortment of aberrations that, together, drive pathogenesis. Regardless of these molecular complexities, there are a set of basic principles that comprise the underlying biology of tumors, known as the hallmarks of cancer[64, 65]. The six hallmarks of cancer include: cells which have an acquired capability to grow self-sufficiently, are insensitive to anti-growth signaling, have limitless replicative potential, sustained angiogenesis, the ability to evade apoptosis, and invade the surrounding tissue as well as extravagate to distal organs (metastasis)[64] (Figure 1). More recently, an emerging body of knowledge indicates two additional hallmarks of cancer: deregulation of cellular metabolism and avoiding immune invasion[65] (Figure 1).

Figure 1: The Hallmarks of Cancer



Adapted from: The Hallmarks of Cancer: The Next Generation
Douglas Hanahan and Robert Weinberg
Cell 2011
https://doi.org/10.1016/j.cell.2011.02.013

The fascinating aspect about each of these traits is that normal, quiescent cells have vast networks of cellular programming, or check points, to stop each of these phenomena from occurring. The ability of a cell to acquire each of these traits is vexingly difficult to imagine, nonetheless the retention of these traits is achieved and observed amongst most, if not all, tumors.

With the advent and technological advances in genetic sequencing techniques, our knowledge of malignant transformation has progressed significantly over the past 20 years, allowing researchers to reliably map the complex network of genetic alterations observed in tumors. Functional studies have revealed that certain molecular pathways are altered at a high frequency across many cancer subtypes, like loss of p53 signaling, a major DNA damage signaling pathway in the cell. While other molecular alterations tend to be tissue specific, like BRCA1 mutations found in breast cancer[116], or even context specific, like the C>A genetic transversions caused by tobacco smoke inhalation[2]. In addition to the alterations observed in the genetic code itself, it has become increasingly apparent that alterations in the molecules responsible for post-translational modifications of the genetic code (epigenetics) are wide spread and play an important role in the initiation, and propagation of cancer as well[153].

1.2 Epigenetics

Epigenetics is defined as 'the study of heritable changes in gene expression that occur independent of changes in the primary DNA sequence[153]. Epigenetics was first defined by C.H. Waddington in 1942 to describe a mechanism that connects a genotype to a phenotype[177]. While this initial observation focused on regulation during developmental processes, it has become increasingly clear that epigenetic dysregulation plays an important role in cancer development as well.

Epigenetic regulation of transcription is inherently intertwined with DNA structure and chromatin organization. The initial process of compacting DNA inside the nucleus is aided by a unique class of DNA binding proteins that wind, loop, and fold DNA into a higher level of structural organization inside the nucleus, called histones[223]. Histones make up the basic building block of the chromosomal structure. DNA is wrapped around histones to form a protein-DNA complex called the nucleosome. Each nucleosome unit consists of two copies of each of the four core histone protein: H2a, H2b, H3, and H4. The resulting histone octamer contains two H2a/b dimers, and one H3/H4 tetramer. DNA is wrapped around the the histone octamer, creating the nucleosome, and each nucleosome is tethered to its neighbor with the help of the peripheral H1 linker histone. From the generation of nucleosomes, to chromosomes, to chromatin, DNA becomes highly condensed and methodically packaged into higher order structures until all the DNA is neatly packaged inside the nucleus, inside the cell.

The ability for the transcriptional machinery to physically access its target DNA is a critical step for initiating transcription, and is an important mechanism used by the cell to regulate gene expression. Specifically, chromatin structure can be modulated into an accessible (open) state to enable recruitment of transcriptional regulatory elements, or vice versa, can be in an inaccessible (closed) state to inhibit recruitment of regulatory elements. Thus, a functional interplay exists between chromatin structure and epigenetic regulators which work together to exert cell-specific and/or context-dependent transcriptional programming.

There are three epigenetic modifications that can alter chromatin structure, modulate chromatin accessibility, and regulate transcription: DNA methylation, histone modifications, or RNA-associated silencing. These three types of epigenetic modifications provide a mechanism for diversity by regulating the genetic information that can be accessed by the cell's transcriptional machinery[153].

1.2.1 DNA Methylation

DNA methylation is a biochemical process where a methyl group is directly added to the DNA, and is often associated to gene silencing. DNA methylation is highly specific, where methyl groups can only be placed between a cytosine and guanine nucleotide linked by a phosphate group, called a CpG site. Interestingly, DNA methylation occurs more frequently at individual CpG sites, but long CG-rich regions >500bp are concentrated in distinct areas of the genome (called CpG islands), and are enriched in gene promoters. Indeed, CpG islands occupy ~60% of all gene promoters in humans, and tend to remain un-methylated[153]. DNA methylation is associated to stable transcriptional silencing, where once the silent state is established, it is generally not revocable. This phenomenon highlights a particularly important role in genomic imprinting and x-inactivation during development[209].

1.2.2 Histone Modifications

Histones can be modified by adding a chemical group, such as methyl, acetyl, phospho, ubiquitin, or small ubiquitin-like modifying (SUMO) groups, to a specific residue of the histone protein[142]. Methylation and acetylation are the two most common histone modifications, and can only be added to the lysine (K) residue of the histone protein. Acetylation involves the transfer of one acetyl group from acetyl coenzyme-A (acetyl co-A) to one lysine residue on the histone protein, and is generally associated with

active transcription. Histone methylation modifications can include the addition of up to three methyl groups onto a single lysine residue, resulting in a mono-, di-, or tri-methylation state. Depending on the location of the lysine residue, the addition of methyl groups can result in transcriptional activation or transcriptional repression. For example, the tri-methylation of lysine 36 on histone H3 (H3K36me3) is an active transcription mark, whereas the tri-methylation of lysine 27 on histone H3 (H3K27me3) is a repressive transcription mark. Unlike DNA methylation, histone modifications are more flexible, where the modification can be added and removed in response to various intracellular or extracellular cues.

1.2.3 RNA Association Silencing

RNA associated silencing is a post-transcriptional modification where small non-coding RNA units (called miRNAs) pair with complementary 3' regions of untranslated target messenger RNA (mRNA) to enact degradation or to inhibit translation[67]. While RNA-associated silencing enacts epigenetic programs to inhibit translation on its own, RNA associated silencing can modulate or be modulated by other epigenetic regulatory mechanisms, highlighting the integration between all transcriptional regulator mechanisms within the cell.

Taken together, the composition of the various biochemical marks across nucleosomes (DNA and histones) comprise the structural landscape of chromatin, and sets the framework for transcriptional programming within the cell.

1.3 Mechanisms of Epigenetic Dysregulation in Cancer

The epigenetic network comprises a series of transcription factors and chromatin modifying enzymes that enact various tissue-specific and context-dependent functions within the cell. Epigenetic alterations can contribute to all facets of tumor development, from potentiation and initiation to tumor progression, and can occur in all facets of epigenetic regulation and chromatin structure, including alterations in DNA methylation, histone modification patterning, miRNA silencing, as well as the proteins responsible for enacting these epigenetic programs. In cancer, the epigenome exhibits vast reorganization of its network to enact oncogenic signaling pathways, activate transcription of oncogenes, and suppress transcription of tumor suppressors. Indeed, while altered epigenome activity contributes to oncogenic signaling, vice-versa, genetic mutations can also contribute to alterations observed in the epigenome. Indeed, ~50% of all tumors exhibit mutations in genes relating to epigenetic regulation and chromatin structure, indicating a functional interplay between genetic and epigenetic alterations that work together to enact oncogenic programming[224]. While the number of epigenetic alterations observed in cancer are far reaching, some recurrent alterations emerge across a wide variety of tumor subtypes, highlighting their importance in cancer development.

1.3.1 DNA methylation

DNA methylation is one of the most extensively studied epigenetic alterations observed in cancer. DNA methylation occurs at genomic regions containing a cytosine and guanine nucleotide linked by a phosphate group (CpG site). While CpG sites occur across the entire genome, they are not evenly

distributed. There are distinct CpG-rich genomic regions, called CpG islands, that are enriched in gene promoters and are generally unmethylated under homeostatic conditions. In many cancers, CpG islands become hypermethylated, which leads to transcriptional silencing of target genomic regions. Functional studies of hyper-methylated CpG islands has uncovered a number of novel tumor suppressor genes that have become epigenetically silenced in cancer, even if not genetically mutated itself[20, 49,70].

One intersting example is the MGMT gene, which encodes for the DNA repair enzyme, O^6-methylguanine-DNA methyltransferase. MGMT removes the naturally recurrung alkyl adducts from the O^6-position of guanine and reconstructs the guanine molecule, which prevents mismatch and errors during downstream DNA replication and transcription[213]. Interestingly, MGMT and its corresponding promoter region is hypermethylated and transcriptionally silenced in approximately 50% of glioblastomas (GBMs)[49]. While it is unclear which mechanism(s) drives this hyper-methylation phenotype in these tumors, it is clear that MGMT is targeted and inactivated by epigenetic alterations, rather than emerging as hotspot genetic lesions. Since its initial discovery, the epigentic status of MGMT has been used as prognostic biomarker in primary glioblastomas[68]. Indeed, MGMT methylation status is an indicator of treatment response to DNA alkylating chemotherapeutic agents, such as temozolomide, which is commonly used for the treatment of glioblastoma[98]. While in normal cells MGMT functions as a tumor suppressor, upon treatment with temozolomide, MGMT can actually become transcriptionally reactivated and tumor-supportive, by counteracting against the chemo-induced DNA damage. Thus, sustained MGMT promoter methylation upon chemotherapeutic treatment with temozolomide correlates with favorable survival in patients with glioblastoma[68]. MGMT is an important DNA repair enzyme which protects the cell from erroneous mutations during replication. The systematic silencing of MGMT in glioblastoma highlights an important epigenetic mechanism used by tumors cells to reduce stability and control over DNA replication. In addition, through epigenetic modifications, GBM cells are able to retain remarkable control over transcription, switching from an inactive to active transcriptional state in response to chemotherapy to evade apoptosis and promote tumor cell survival[98].

MGMT is just one example of tumor-driven DNA methylation and transcriptional silencing, however there are many other targets of DNA hyper-methylation observed in cancer. Some examples include the cyclin dependent kinase inhibitors (CDKN2a and CDKN2b), the Ras-associated domain family 1 isoform A (RASSF1A), androgen receptor (AR), and adenomatous polyposis coli (APC), among others[20, 38, 82, 132]. Each of these genes have important tumor suppressive functions, being involved in DNA damage response, cell cycle regulation, corticosteroid signaling, and cell proliferation and survival. In addition, these genes are common targets of DNA-methylation mediated transcriptional silencing in cancer.

Protines involved in the generation and establishment of DNA methylation patterning are also frequent targets of genetic mutations in tumor cells. For example, isocitrate dehydrogenase 1/2 (*IDH1* and *IDH2*) mutations are frequently observed in glioblastoma, and the presence of a mutation in *IDH* is correlated to global CpG-island hyper-methylation and subsequent transcriptional silencing of tumor suppressor genes. Specifically, mutations in the *IDH* gene leads to a novel function of the IDH protein, namely the production and accumulation of D-2-hydroxyglutatarate (2HG). The accumulation of 2HG leads to

the competitive inhibition of α-ketoglutarate-dependent dioxygenases, and subsequent accumulation of DNA and histone methylation in *IDH*-mutant cells[178]. The mutations in *IDH* observed in glioblastoma is an interesting example of the functional interplay between genetics and epigenetics, where a genetic mutation results in a novel function of the encoded protein, which gives rise to epigenetic alterations, and aberrant metabolic signaling.

In addition, mutations in the enzymes responsible for catalyzing the transfer of methyl groups onto the DNA are also frequently observed in cancer. In mammals, there are three DNA methyltransferase enzymes, DNMT1, DNMT3a, and DNMT3b. Oncogenic mutations in DNMTs result in genome-wide alterations to the DNA methylation landscape in many tumor subtypes, particularly hematologic malignancies, including acute myeloid leukemia (AML)[102], myelodysplastic syndrome (MDS)[181], t-cell lymphoma[36], and colon cancer[84].

Taken together, DNA methylation plays an important role in tumorigenesis, and is altered in a variety of different ways to enact oncogenic programming in tumor cells. On one side, the promoter regions of tumor suppressor genes can become hyper-methylated to shut down tumor-suppressive activity within tumor cells. While on the other side, tumor cells can exhibit genetic mutations within the DNA methylation machinery that lead to global changes in methylation landscape and wide-spread changes in transcriptional programming of the cell.

1.3.2 Histone Modifications

In addition to modifications observed at the DNA Methylation level, a number of alterations in histone remodeling factors have been observed in cancer as well. Histones can be modified by the addition or removal of various chemical groups, such as acetyl and methyl groups, which are the two most common histone modifying groups found in human cells. Acetyl groups are added or removed from histones by histone acetyltransferases (HATs) and deacetylases (HDACs), respectively. Methyl groups are added or removed from histones by histone methyltransferases (HMTs) and demethylases (HDMs), respectively. Alterations in all four classes of enzymes have been observed widely across many cancer subtypes, and herein, a few examples of cancer-associated histone modifications are discussed.

First, the AML1-ETO chimeric fusion protein generated by the t(8;21) translocation is one of the most common alterations observed in acute myeloid leukemia (AML). Interestingly, this fusion generates a novel DNA interaction site for the HAT enzyme, p300, leading to the active transcription of the AML1-ETO fusion protein, which is important for leukemic cell transformation in this context[183]. It has been shown that the HAT activity of p300 is critical for the transcriptional activating function of AML1-ETO, and loss of p300 leads to a decrease in tumor cell growth *in vitro*[183]. In this context, the p-300 mediated histone acetylation is required for the transcriptional activity of the AML1-ETO oncogenic fusion protein, both of which are generated by the t(8;21) translocation and requierd for leukemic cell transformation. This particular example outlines an interesting mechanism of epigenetic hijaking because the precise site of the t(8;21) translocation not only generates the genetic code for the synthesis of an oncogenic fusion protein, but also ensures its transcription by creating a new site for transcriptional activation through p300 binding and subsequent acetylation of the histone protein.

Histone modifications can also occur by directly mutating the histone modifying machinery. For example, *EZH2* is the histone methyltransferase (HMT) of the polycomb repressive complex 2 (PRC2), which is responsible for transcriptional silencing through the mono- di- and tri- methylation of lysine 27 on histone-H3 (H3K27). A single mutation in *EZH2* has the potential to alter the global methyl landscape of H3K27, which results in modifications to the transcriptional landscape observed in these cells. Mutations in *EZH2* provide an example of how one single genetic mutation can have widespread effects on the chemical modifications of histones and thus transcrptional programming within the cell. However, the types of mutations in *EZH2* observed in cancer can be highly context and tissue dependent. For example, enzymatic gain-of-function *EZH2* mutations are often observed in lymphoma, which results in a global increase in H3K27me3 and subsequent transcriptional silencing[119]. In solid tumors, *EZH2* is often overexpressed or amplified, also resulting in a global increase in H3K27me3[27]. However, in myeloid malignancies, enzymatic loss-of-function mutations are observed, indicating cell-type specific oncogenic activity of the EZH2 enzyme[173]. In addition to mutations observed in *EZH2*, alterations in other PRC2 complex proteins are also frequently observed in cancer, and have the ability to modulate the HMT activity of EZH2 and the PRC2 complex. The PRC2 core subunit proteins, *EED* and *SUZ12*, are targets of inactivating mutations or deletions in T-cell acute lymphoblastic leukemia (T-ALL), and are associated to the activation of Notch signaling due to loss of H3K27me3[124]. Taken together, alterations to the HMT EZH2 enzyme, whether activating or silencing, result in wide-spread effects on the histone methylation patterning in the cell, leading to genome-wide transcriptional re-programming that contributes to oncogenic programming in a tissue-specific manner.

The variety of alterations in histone modifying enzymes observed in cancer highlights a particularly important role in epigenetic transcriptional regulation, where alterations can lead to both activation or repression of transcription in a context-dependent manner, seeminly designed to enhance the fitness and survival of these cells. The extent of histone modifications observed in cancer is widespread, and the exact role of each alteration is not completely understood. Nonetheless, a growing body of knowledge on the mechanisms guiding these alterations continues to expand our understanding of how modifications to histones work to exert oncogenic signaling and tumor-specific functions.

1.3.3 Chromatin Remodeling and Positioning

Nucleosomes play an important part in transcription, and a number of protein subunits are responsible for maintaining nucleosome structure in an ATP-dependent manner[191]. Mutations in the nucleosome remodeling machinery are often observed in cancer, pinpointing an important link between chromatin remodeling and oncogenic programming. For example, the SWI/SNF (BAF) complex is a large nucleosome remodeling complex with multiple varieties of subunit assemblies, and is silenced in many cancer subtypes. The prevailing hypothesis is that while many individual SWI/SNF subunit complex proteins have tumor suppressive activity, the systematic targeted silencing of the multiple SWI/SNF complex subunits leads to epigenetic and genomic instability by changing nucleosome occupancy patterning throughout the nucleus, thereby altering tumor suppressive programming[144].

1.4 Targeting the Epigenome in Cancer

The epigenome plays a hierarchical role in regulating transcription in the cell, where a mutation in a single epigenetic regulator can lead to alterations in many cell signaling pathways. The extent of epigenetic alterations observed in cancer lends an attractive path for the development of targeted therapeutic approaches. Indeed, a number of inhibitors targeting epigenetic regulators have shown promise in the clinic, and herein, a few clinical examples are outlined.

DNA methyltransferase (DNMT) inhibitors were the first clinically approved class of epi-therapeutics used for the treatment of cancer. Azacitidine and 5-Aza-2-deoxycytidine (Decitabine) are two compounds that act as competitive inhibitors of DNMTs in the cell by binding the cytidine nucleotide and blocking the deposition of methyl groups onto target CpG sites, thus inducing de-methylation and reactivation of previously silenced genes[149]. Aza-nucleosides have shown efficacy in treating various hematologic malignancies, including acute myeloid leukemia (AML), t-cell leukemia, and myelodysplastic syndrome (MDS)[56, 57]. However, while still effective in treating many hematologic malignancies in the clinic, aza-nucleosides show high cytotoxicity and poor chemical stability *in vivo*, limiting their potential to be effective in treating solid tumors[60]. More recently, a chemical variant of decitabine has emerged (CC-486, Celgene), which allows for oral ingestion and increased half-life of the chemical compound *in vivo*, and is currently in clinical trial for the treatment of breast and lung carcinomas (clinical trial ID: NCT02223052, NCT02250326, NCT02546986).

One reason for the high cytotoxicity observed with aza-nucleosides is that they target all DNA that is actively replicating in the body, and do not specifically target tumor cells. To specifically target tumor cells, new compounds have been developed to target the upstream mutations responsible for the hyper-DNA methylation phenotype observed in tumor cells specifically. For example, a number of IDH-mutant specific inhibitors (AG-120, AG-221, AG-881, BAY1436032, IDH305) have been developed to inhibit the mutant IDH protein leading to the inhibition of D-2-hydroxyglutatarate (2HG) production and subsequent reversal of the hyper DNA and histone methylation observed in tumor cells[178]. Treatment with IDH-mutant inhibitors restores the DNA and histone methyl-landscape and induces cellular differentiation in hematopoietic malignancies[165, 194]. Indeed, pre-clinical and clinical studies of IDH-m inhibitors have shown significant response rates and survival benefits for patients with IDH-m tumors (Clinical trial ID: NCT02273739, NCT02073994, NCT02481154, NCT02746081, NCT02977689). In addition, these orally available small molecule inhibitors are highly selective with relatively few side-effects arising with treatment[166]. Treatment with small molecular inhibitors that target IDH-mutations outline one specific example of selectively targeting genetic mutations to reverse DNA methylation alterations that contribute to tumor development and progression in this context.

In addition to targeting alterations in aberrant DNA methylation patterning, a number of inhibitors targeting histone modifications have been developed as well. In a similar mechanism as DNA methylation inhibitors, two classes of inhibitors have been developed: those targeting histones globally, and those targeting the upstream mutations responsible for enacting hyperactivity of histone modifications. Both classes of small molecule inhibitors have been applied in the clinic for the treatment of various cancer

subtypes.

Small molecule inhibitors targeting histone deacetylases (HDACs) increase the accumulation of acetylated histones leading to an open chromatin state. The result of hyperacetylation of histones leads to both the activation and repression of gene expression, indicating a diverse regulatory role of acetylation in nuclear receptor signaling functions, that eventually leads to cell cycle arrest and apoptosis[147]. Two HDAC inhibitors have been approved for the treatment of T-cell lymphoma, vorinostat and romidepsin, showing significant improvements in overall survival in patients[8, 31, 46, 127, 190]. Further pre-clinical studies have broadened the application for HDAC inhibitors to a number of additional tumor subtypes, resulting in the initiation of clinical trials assessing the efficacy of HDAC inhibitors in other hematologic and solid tumors, including mantle cell lymphoma, melanoma, breast, and lung cancer (clinical trial ID: NCT02836548, NCT02151721, NCT03742245).

Inhibitors targeting recurrent mutations in the histone modifying machinery have also been applied to the clinical setting. One clinical example applies to mutations frequently observed in the histone methyltransferase, *EZH2*. Mutations in *EZH2* are frequently observed across a variety of tumor subtypes, but are particularly prevalent in non-Hodgkin lymphomas (NHL), where *EZH2* mutations are identified in approximately 20-30% of patients[11, 119]. Mutations in *EZH2* alter the HMT activity of PCR2, resulting in hyper-methylation of target histones and subsequent transcriptional silencing[159]. A number of inhibitors targeting *EZH2* mutations have been tested in the clinical setting for the treatment of *EZH2*-m lymphoma, which function to reduce the HMT activity of the EZH2 protein leading to the reactivation of previously silenced target genes. While a number of chemical inhibitors of EZH2 have been developed over the years, the most recent EZH2 inhibitors act as selective inhibitors of S-adenosyl methionine (SAM) pocket of the catalytic set domain of EZH2, where SAM acts at the universal methyl donor for HMTs. Indeed, while the early EZH2 inhibitors did not show clinical benefit and clinical trials were halted (GSK2816126, GlaxoSmithKline; NCT02082977), next generation EZH2 inhibitors (EPZ-6438, Epizyme and CPI-1205, Constellation Pharmaceuticals, Inc.) have shown promise in early phase clinical trials for the treatment of NHL, leukemia, and other solid tumors[78] (NCT02395601).

Epigenetic therapies offer an intriguing method for targeting the multi-faceted, hierarchical transcriptional machinery in cancer. Indeed, a number of inhibitors have been developed to target the epigenetic machinery in a variety of different ways, and have shown promise in the pre-clinical and early clinical trials of many cancer subtypes. However, epigenetic dysregulation in cancer is an extremely complex process whose consequences can vary dramatically depending on the cell or tissue of origin. Functional studies that incorporate analyses of epigenetic regulation, chromatin architecture, and transcriptional programming will continue to unravel the complex mechanisms of oncogenic programming and help to identify exploitative vulnerabilities for therapeutic intervention.

1.5 Aim and Scope of Thesis

Epigentic regulation is an important mechanism that allows the cell to maintain control over which genes are actively transcribed at any given time, thereby altering the cell's phenotype without alterting the genetic code itself. Epigenetic regulation of transcription occurs as a result of various intracellular and extracellular cues, where the resulting cellular phenotype can be inherited from generation to generation and ultimately contributes to the remarkable heterogeneity observed in humans and other higher order organisms.

Cancer cells exhibit extensive dysregulation of epigenetic machinery and epigenetic control of transcription, allowing the tumor cell to alter various cell signaling pathways, to activate oncogenes, and silence tumor suppressor genes in one fell swoop. In tumors, a functional interplay exists between epigenetic alterations and genetic mutations that work together to alter chromatin structure and enact oncogenic signaling. The extent of epigenetic reprogramming and types of epigenetic alterations observed in cancer can be highly context dependent and tissue-specific. Therefore functional studies aimed to improve our understanding of the interdependencies between epigenetic alterations and genetic mutations and their contrubtion to cancer development and progression in different contexts will provide valuable insight into the various mechanisms of oncogenic programning observed in human cancers.

In this thesis, we explore the mechanisms of epigenetic dysregulation in cancer in two separate contexts.

Chapter 2: EZH2 oncogenic mutations drive epigenetic, transcriptional, and structural changes within chromatin domains

In the first study, we explore how gain-of-function mutations in the PRC2 histone methyltransferase, enhancer of zeste homologue-2 (*EZH2*), alter the structure and function of topologically associating chromatin domains in non-Hodkin Lymphoma. Approximately 20-30% of diffuse-large b-cell (DLBC) and follicular lymphomas harbor a somatic point mutation in *EZH2*, called $EZH2^{Y646X}$. $EZH2^{Y646X}$ mutations lead to a genome wide increase in H3K27me3, which is associated to transcriptional silencing. Given the association between transcriptional silencing and chromatin structure, we sought to explore these associations in the context of $EZH2^{Y646X}$ in NHL. In this study, we combined transcriptomics and epigenetics with analyses of chromatin structure, to uncover how the $EZH2^{Y646X}$ epigenetic mutations drive global changes in transcription to drive tumorigenesis by modulating chromatin interactions and chromatin three-dimensional structure.

Chapter 3: Dissecting epigenetic, transcriptomic, and metabolic signaling dependencies using a novel model of central nervous system primitive neural ectodermal tumors (CNS-PNETs)

In our second study, we dissect the epigenetic, transcriptomic, and metabolic signaling dependencies of central nervous system primitive neural ectodermal tumors (CNS-PNETs). CNS-PNETs are a particularly aggressive subtype of embryonal CNS-tumor that exhibit distinct genetic and epigenetic features, and have a significantly worse clinical outcome compared to other pediatric CNS tumors. Furthermore, diagnosis and patient stratification of CNS-PNETs is limited to descriptive MRI analyses, which solely provide insight into the anatomic location and size of these tumors, but lend little information to aid in diagnosis and treatment response.

In the absense of distinct driving genetic mutations, it has been hypothesized that CNS-PNETs are instead driven by aberrant DNA methylation linked to complications during neural development[208]. However, a lack of *in vivo* models for studying CNS-PNETs has challenged the opportunity to dissect the genetic and epigenetic variables that underlie the development of these tumors.

Due to the lack of models for studying CNS-PNET pathogensis, our first aim was to develop a novel *in vivo* model of CNS-PNET tumors using neural progenitors cells dervied from human induced pluripotent stem (hIPS) cells, called the CNS-NPC model. In addition, we utilized the only CNS-PNET patient-derived cell line as a second model of CNS-PNET pathogenesis in our study (called PFSK-1). We performed a comprehensive characterization of CNS-NPC and PFSK-1 tumor models, including analyses of tumor growth, histology, immunochemistry, gene expression and DNA methylation profiling, and confirmed that that the epigenetic and molecular features of distinct subtypes of CNS-PNET tumors are recapitulated by the CNS-NPC and PFSK-1 models. Finally, we utilized the CNS-NPC and PFSK-1 tumor models for pilot studies to assess the efficacy of metabolic imaging platforms in the diagnosis of CNS-PNETs. By employing both *ex vivo* and *in vivo* metabolic analyses of CNS-NPC and PFSK-1 tumors, we identified a number of metabolic alterations in CNS-PNET tumors that are detectable using ^{1}H nuclear magnetic resonance imaging, as well as ^{13}C hyperpolarized magnetic resonance imaging, including glycine, taurine, and choline. In addition, we identified two metabolic markers, lactate and myo-inositol, which are associated to tumor growth, and varying stages of tumor aggressiveness observed between the two models.

Chapter 2: EZH2 oncogenic mutations drive epigenetic, transcriptional, and structural changes within chromatin domains

2.1 Introduction

2.1.1 Lymphoma

Lymphoma accounts for approximately 5% of all cancers,and over 50% of all blood-specific cancers globally (WHO). Lymphoma is a cancer that develops specifically in the lymphocytes (i.e. b-cells, t-cells, and natural killer cells), which collectively form the adaptive immune system. Lymphoma can originate in all types of lymphocytes and are categorized as such into several classes and subclasses.

In the context of b-cell lymphoma, diagnosis is classified into two main sub-types: Hodgkin lymphoma (HL) and non-Hodgkin lymphoma (NHL). The Reed-Sternberg cell, a cell type that is present specifically in HL, is primarily used to distinguish HL from NHL. NHL accounts for all other lymphomas (85-90% of all cases), and is further classified based on the pathological features observed during b-cell differentiation. Diffuse large b-cell lymphoma (DLBCL) is one of the most common and aggressive forms of NHL, accounting for approximately 40% of all NHL cases. Other less common subtypes of NHL include follicular lymphoma, mantle cell lymphoma, lymphoplasmacytic lymphoma, marginal zone lymphoma, and small-cell lymphocytic lymphoma (SLL) (WHO).

Recent molecular analyses of DLBC lymphoma have revealed distinctions between tumors that contribute to varying responses to treatment. In the clinic, first line treatment for DLBCL primarily consists of anti-CD20 antibody (Rituximab) and the generic chemotherapies: Cyclophosphamide (Cytoxan), Doxorubicin Hydrochloride (Adriamycin), Vincristine Sulfate (Oncovin), and Prednisone (collectively known as R-CHOP combination)[197]. Clinical response to R-CHOP is highly heterogeneous, with only about 40% of patients responding to therapy, underlining a pressing need to identify better ways to stratify patients for treatment optimization[203]. In early 2000, molecular analyses revealed distinct subtypes of DLBCL: germinal center b cell-like (GCB) and activated b cell-like (ABC), which are representative of b-cell tumors that originate from different stages during b cell development[203, 218]. These different signatures are not absolute, meaning that a certain degree of heterogeneity exists across patient samples. Nonetheless, these signatures help to provide a better understanding of the unique signaling dependencies in DLBCL and to stratify patients, with regard to treatment response, progression free survival, and overall survival. Germinal center b cell-like (GCB) tumors have transcriptional signatures characteristic of normal germinal center b cells, including expression of genes like *BCL6*, a well established marker for germinal center b cells, as well as the cell surface proteins, CD10 and

CD38[203]. Activated b cell-like (ABC) tumors have a significantly worse clinical outcome than those of the GCB subtype. ABC tumors show hallmark transcriptional signatures of the mitogenically activated peripheral b cells, up-regulating expression of genes like *BCL2*, a critical component of the b cell receptor signaling pathway, and *IRF4*, an important component of b-cell proliferation following antigen receptor stimulation[203].

As one might expect from such variable expression signatures, GCB and ABC DLBC lymphomas also tend to have distinct genomic alterations and signaling dependencies. GCB lymphomas exhibit de-regulation of the phosphatase and tensin homologue (PTEN), phosphatidylinositol 3-kinase (PI3K), and the janus kinase/signal transducers and activators of transcription (JAK/STAT) signaling pathways[175]. In addition, GCB lymphomas exhibit recurrent genomic alterations in chromatin regulators or modifiers like *KMT2D (MLL2), CREBBP and EZH2*[85, 205]. ABC lymphomas tend to exhibit recurrent mutations in genes involved in b cell receptor signaling, including *CD79a/b, CARD11, BCL10, BTK, IRF4*, and *IRF8*[85, 155]. As a result of chronic b cell receptor signaling activation, ABC tumors have constitutive activation in the NF-kB pathway. Since early 2000, when the distinctions between various DLBC lymphomas subtypes were first established, there has been an enormous effort to better characterize these molecular subtypes (ABC and GCB) to inform and implement improvements for clinical treatment[30, 197].

2.1.2 Polycomb Group Proteins

Polycomb group proteins (PcG) are involved in epigenetically-mediated transcriptional silencing through the post-translational modification of target histone proteins. Specifically, the polcomb group proteins methylate histone H3 at lysine 27 (H3K27 methylation). The PcG proteins were initially discovered in fruit flies (*Drosophila melanogaster*), where they were identified as transcriptionally repressive regulators of the Hox-family genes[45, 101]. Hox gene expression patterns are established early in embryonic development, and are responsible for cell differentiation patterning along the anterior-posterior axis of the fly[77]. However, the early transcription factors of the Hox genes are short lived, decaying rapidly in the initial stages of development, but the transcriptional status of these genes are maintained throughout development by two protein groups: PcG for maintaining a silent state, and trithorax group (TrxG) for maintaining an active state[86, 137]. Since their initial discovery, PcG proteins have been shown to play critical roles in maintaining identity of stem, progenitor and differentiated cells far beyond simply the regulation of the Hox-family genes[34]. PcG proteins have been identified in a number of different species, from plants to animals, including worms, mice, and humans[16, 99, 157, 189]. Both structural and functional analyses of the PcG protein complexes reveal remarkable similarity across several species, underlining its important role throughout evolution[75, 100]. Despite such similarities, there are also some clear divergences. For example, some plants species can contain as many as 12 different PcG complex variants all with unique functions, while humans have only two PcG core complexes, PRC1 and PRC2, which seem to exhibit highly overlapping function[69].

In mammals, the PRC complexes have been found to regulate many cellular processes through transcriptional silencing of genes involved in cell plasticity and differentiation, cell fate determination, cell cycle regulation, X-inactivation, and senescence[182, 204]. PRC1 functions to catalyze the monoubiquitylation of lysine 119 of histone H2A (H2A119ub) and is thought to mediate transcriptional

silencing through the physical compaction of chromatin mediated by the HPC proteins (CBX2, CBX4, CBX7, CBX8)[61]. PRC2 mediates transcriptional silencing through the di- and tri-methylation of histone3 lysine 27 (H3K27), which is carried out through the histone methyltransferase activity of a core subunit protein, EZH1/EZH2[1]. During early development, PRC proteins play an important role in the maintenance and regulation of stem cell identity, binding up to 20% of gene promoters in embryonic stem (ES) cells[17, 99]. For many early cell-fate decisions, the stable PcG epigenetic footprint is maintained throughout multiple cell divisions, where existing histones maintain their epigenetic marks and newly synthesized histones gradually acquire their post-translational modifications during the cell cycle[202].

While the H3K27 methylation fingerprint exists to maintain transcriptional repression throughout multiple cell divisions, it has also been shown that PcG-mediated transcriptional silencing can be a dynamic process as well. This dynamic action between inactive and active states can be carried out by a class of histone demethylases, including JMJD3 (KDM6B) and UTX (KDM6A), which can remove PcG protein-dependent histone modifications enabling reactivation of previously silenced genes[151, 167]. Another example of transcription state switching is observed during development. Some genes that are initially silenced by PRC can become activated, and vice versa, through the establishment of bivalent domains. For example, in early ES cells, many H3K27me3 regions also contain the active histone mark, H3K4me3. Many targets that are silenced early in development are poised for activation at a later stage of differentiation, where at the right time, the target loses the H3K27me3 mark and gains the H3K4me3 mark, thus becoming transcriptionally activated. Vice versa, a poised repressive bivalent domain will eventually lose the active H3K4me3 mark and gain the repressive H3K27me3 mark. The choice between a maintained versus dynamic silent state depends on the PRC-target region, as well as other factors (like histone demethylases), which guide such cell-specific and/or context-specific decisions[143]. In summary, the PRC complexes have an important role in transcriptional regulation during development and throughout the entire life cycle of a cell that is carried on through generations. In specific contexts, PRC-mediated transcriptional regulation can also be dynamic, allowing the cell to adjust to both intra- and extra-cellular cues.

2.1.3 PRC Binding to Chromatin

The PRC complexes function to maintain transcriptional silencing by depositing methyl and ubiquitin groups onto target histones, thereby inhibiting transcriptional machinery from accessing DNA. The ability for the PRC complexes to bind to chromatin is a critical step in exerting its epigenetic program.

The mechanisms by which PRC2 is recruited to target loci in mammals is unclear. Unlike the D. melanogaster PRC2 homologue, where PRC2 is recruited to a specific DNA sequence (called a polycomb-responsive element, or PRE), only a few human PRE's have been identified to date. Yin Yang 1 (*YY1*) was previously identified as a PRC2 binding protein in certain tissue and cell types[5, 24], however *YY1* has also been implicated in a number of other processes, indicating that it does not function exclusively as a mammalian PRE. More recently, a more general PRC2 target sequence was identified in mammalian genomes, whereby PRC2 binds to large GC rich elements contained in CpG islands called G-quadruplex sequences [G4N1-7]4 [115]. However, this model integrated bacterial artificial chromosomes (BACs) from the Escherichia coli genome to functionally demonstrate this binding

capacity, but these naturally occurring elements in mammals have yet to be functionally annotated. Nonetheless, there are a number of possible mechanisms of PRC2 recruitment. First, DNA binding motifs (PRE), such as G-quadruplex sequences, may indeed preferentially bind PRC2, however further studies are required to determine how specific these sequences are for PRC2 and their exact functional relevance in mammals. In addition, there may be cell-specific or gene-specific transcription factors or non-coding RNA that act as PRC2 recruiters[185], but, as in the case of *YY1*, may be context or tissue specific. Finally, histone variants within a cell (whether structural or biochemical) may also play a role in PRC2 recruitment and binding to chromatin.

Genome-wide mapping studies of PRC1 and PRC2 in human cells confirmed that the two complexes tend to co-localize on the genome together with H3K27me3[17, 99]. Studies in mouse and human embryonic cells have shown that PRC1 binds to PRC2 methyl regions of H3K27, meaning that PRC2 could play a role in facilitating PRC1 binding[184]. Given this binding preference, it has been suggested the PRC2 establishes and maintains transcriptional silencing and chromatin compaction through PRC1[89, 207]. However, this mechanism has not been well established in the context of human cells. First, a number of reports have shown that a variety of different PRC1 complex subtypes exist, and these subtypes may each have different functions within the cell[59]. In addition, it has been shown that some genomic regions are only occupied by PRC2, and do not additionally contain PRC1[94]. And finally, while human PRE's have yet to be identified, a report in 2009 identified the first mammalian PRE in mouse cells (called kr-PRE). Interestingly, it was shown that this kr-PRE is PRC1-specific and does not bind PRC2, suggesting that recruitment of PRC1 is not completely dependent on PRC2, at least in this context[158]. Taken together, it is likely that in mammals (and humans in particular), there are both overlapping and distinct roles of PRC1 and PRC2 binding to mediate transcriptional repression, and further research in the field will continue to unravel these distinctions.

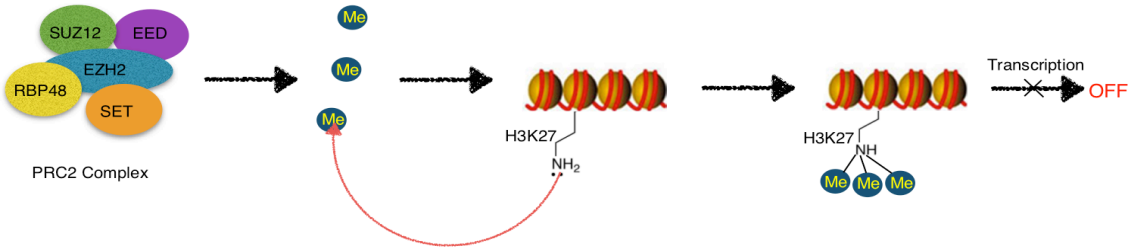
2.1.4 The PRC2 Complex

The PRC2 complex is comprised of four core components required for its activity: EZH1 or 2 (enhancer of zeste homologue 1/2), SUZ12 (suppressor of zest 12), EED1-4 (embryonic ectoderm development transcript variants 1-4), and RbAp46/48 (retinoblastoma-binding protein p46/48; also known as RBBP7/4). The primary role of PRC2 is the di- and tri- methylation of lysine 27 on histone H3 (H3K27), which is carried out through the catalytic activity of EZH1/2. While PRC2-EZH1 and PRC2-EZH2 both preferentially utilize H3K27me1 for substrate binding and have many shared target regions, the two complexes can function differently in the cell. The most apparent evidence for this is exhibited by the histone methyltransferase (HMT) activity of the two proteins, where EZH1 has approximately 20-fold lower HMT activity than that of EZH2[108]. In a genetic knockdown experiment of *EZH2 in vitro*, global levels of H3K27me2 and H3K27me3 were significantly reduced, and H3K27me1 was slightly elevated. However, in a similar experiment with *EZH1*, no significant changes on H3K27me1, 2 or 3 were observed, indicating highly varied roles for PRC2-EZH1 and PRC2-EZH2 in mammalian cells[108].

While the catalytic activity lies with the EZH proteins, the non-enzymatic subunits, SUZ12 and EED, are also required for the histone methyltransferase (HMT) activity of EZH1/2. Specifically, EED aids in PRC2 recruitment to H3K27me3 and also helps to stimulate the HMTase activity of EZH1/2. SUZ12

provides structural integrity to the PRC2 complex and EZH1/2 itself[154, 163]. While RbAp46/48 is not required for the HMTase activity of EZH1/2, the protein subunit acts as a histone chaperone, binding H3-H4 heterochromatin to help mediate PRC2 recruitment and maintain chromatin stability[154].

Figure 2: The PRC2 Complex



PRC2 methylation of Histone 3 Lysine 27 and it's effect on transcription
Adapted from Olaneli
CC Attribution-Share Alike 4.0 International license

In addition to the four PRC2 subunits, there are also transient proteins that have been shown to interact with the PRC2 complex. Some of the cofactors have a known function with respect to PRC2, including JARID2, AEBP2, and the polycomb-like genes (PCL) 1-3 (PHF1, MTF2, and PHF19, respectively). PCL1-3 and JARID2 aid recruitment of PRC2 to chromatin, particularly during cell lineage commitment, most often described in mouse ES cells[96, 150, 200]. Specifically, JARID2 contains a DNA binding domain which stabilizes binding of PRC2 to chromatin, and has been shown to play a role in the catalytic activity with both activating and repressive properties[133, 150]. PCL1-3 have highly specific recognition sites for H3K36me3, which work to inhibit PRC2-mediated H3K27 methylation in recognized chromatin[120, 141]. AEBP2 is a zinc finger protein that binds to multiple components of the PRC2 complex to enhance its enzymatic activity. AEBP2 also contains a DNA binding domain, which may aid in PRC2 recruitment to chromatin[112]. Additional proteins have also been shown to interact with PRC2, including HDACs, SIRT1, and DNMTs, yet their role or function remain unclear[110, 180]. These transient proteins are not always found to be associated to the PRC2 complex, suggesting an important role for PRC2 to have the flexibility in modulating its binding and/or enzymatic activity in the cell. Given the number of variants between the core subunit proteins as well as the number of transiently interacting proteins, the PRC2 complex comes in many different forms and flavors, all of which contribute to PRC2 functional diversity within the cell.

2.1.5 EZH2 mutations in lymphoma

Both loss and gain of function mutations in the PRC2 core complex proteins have been identified in many cancer types, and are generally associated with poor prognosis. For the purpose of this study, we focus specifically on mutations that arise in the histone methyltransferase of the PRC2 complex, *EZH2*. Mutations in *EZH2* have been identified in a number of different tumor types, including non-hodgkin lymphoma (NHL), follicular lymphoma[119], melanoma[161], ewing sarcoma[171], and many solid

tumors including prostate and breast cancers[27] (www.cbioportal.org/genie). The types of mutations observed can vary dramatically, with specific mutations occurring more frequently within specific tissue or cancer subtypes. Furthermore, the phenotype and functional consequence of such mutations can vary dramatically. Many solid tumors, including breast, lung, and prostate cancers exhibit *EZH2* overexpression or amplifications. In hematological malignancies, *EZH2* seems to have opposing roles depending on the cell of origin. In b-cell lymphomas, *EZH2* has an oncogenic function[119], but in myeloid malignancies *EZH2* has a tumor suppressive function[173], indicating diverse roles of action of EZH2.

In developing b-cells, *EZH2* is required for germinal center (GC) formation. Specifically, in GC b cells, *EZH2* expression helps to speed up the G1 to S phase transition to facilitate rapid cell division and clonal expansion during b cell differentiation[9]. In the context of b cell lymphoma, a specific somatic point mutation emerges as a recurrent alteration in the GCB subtype of DLBC lymphomas (~20-30%) and follicular lymphomas (~5-10%), where the tyrosine (Y) residue at position 646 is changed to phenylalanine (F) or asparagine (N), and is more commonly referred to as $EZH2^{Y646X}$ [11, 119]. In it's initial discovery, $EZH2^{Y646X}$ was described as a loss of function mutation[119], however, it was quickly determined that the $EZH2^{Y646X}$ mutations observed in tumors is in fact a gain-of-function heterozygous mutation. Interestingly, $EZH2^{Y646X}$ homozygous mutations are embryonic lethal. Instead, $EZH2^{Y646X}$ requires the wild-type *EZH2* allele to modulate substrate specificity, which leads to a global increase in H3K27me3[159]. Furthermore, in a CD19-dependent Cre-conditional $EZH2^{Y646X}$ mouse model, $EZH2^{Y646X}$ was sufficient to induce lymphoma and melanoma development *in vivo*[161]. Thus, $EZH2^{Y646X}$ mutations function to promote oncogenic programming of DLBCL and follicular lymphoma.

2.1.6 Chromatin organization

The advent of techniques aimed to map the three-dimensional conformation of chromatin, called chromatin conformation capture, has allowed researchers to begin to unravel the importance of the hierarchical structure of chromatin in regulating fundamental signaling processes inside the cell. Briefly, chromatin conformation capture involves the crosslinking of histones with formaldehyde, which creates bonds that "freeze" interacting genomic loci in space and time. Restriction enzymes are then used fragment DNA, and are used as recognition sites for subsequent mapping to the human genome in downstream processing. Following restriction enzyme digestion, proximity ligation allows for the annealing of two DNA fragments that are close to one another in 3D space but not necessarily connected on the linear genomic template. The resulting fragments of interacting DNA are then prepared for high-throughput sequencing. The resulting output allows one to quantify the number of genomic interactions that occur in 3D space, and can be mapped to analyze the spatial organization of chromatin inside the cell at multiple levels of hierarchical structure.

Studies of three-dimensional chromatin conformation have identified a hierarchical pattern of chromatin looping and folding inside the nucleus of eukaryotic cells[13, 43, 52, 41]. Hierarchical structural features can include loops made of relatively short sequences (~1-10kb), medium size domains of densely interacting chromatin (~100kb-1Mb), also known as topologically associating domains (TADs), and large chromatin compartments (> 1Mb), which segregate genomic regions that are either actively transcribed or silenced[105].

Most comparative studies of chromatin architecture have focused on developmental processes, including cell differentiation and cell fate commitment[14,15, 92, 192]. Studies of genome architecture in relation to human disease have begun to gain attention in recent years, particularly in the interest of improving our understanding of the functional consequences of cancer-associated alterations. In one example of t-cell acute lymphocytic leukemia (t-ALL), a recurrent deletion of chromosome 1p33 leads to the creation of a new super-enhancer (marked by H3K27ac) that hijack transcription and activate expression of TAL1, an important oncogenic transcription factor in t-ALL[109]. Another study of gliomas demonstrated a novel role for the gain-of-function isocitrate dehydrogenase (*IDH*) mutation in chromatin remodeling by hypermethylation of the CTCF protein binding sites. This epigenetic alteration leads to the reduction of CTCF biding and loss of TAD boundaries, leading to aberrant enhancer activity and transcriptional activation of the platelet derived growth factor receptor alpha (PDGFRA) glioma oncogene[51]. Both studies present unique methods of oncogenic transcriptional reprograming by modifying chromatin organization, and further studies combining 3D architecture with transcriptomics and epigenetics will continue to unravel this functional interplay.

2.2 Rationale for Project

It has been speculated that, in the context of *EZH2*^{Y646X} mutations, such a dramatic global increase in H3K27me3 may lead to a similarly vast reorganization of chromatin within the nucleus, and the advent of high-throughput sequencing techniques combined with chromatin conformation capture has enabled further research on this topic. TADs have been proposed as candidate functional elements that work to regulate gene expression during development and other epigenetic differentiation programs[52, 41, 91]. Indeed, it has been shown that TADs are preferentially enriched for active or repressive transcription marks, like H3K36me3 or H3K27me3, respectively[97, 122, 152]. Given that *EZH2*^{Y646X} has been shown to increase H3K27me3 and promote transcriptional repression genome wide[219], we wondered whether oncogenic *EZH2* mutations lead to structural and/or functional modifications of chromatin. To address this question, we investigated epigenetic, transcriptional, and structural changes in *EZH2* wild-type and mutated lymphoma cells in the context of their chromatin architecture.

2.3 Main Findings

EZH2 oncogenic mutations drive epigenetic, transcriptional, and structural changes within chromatin domains

Maria Donaldson Collier*, Stephanie Sungalee*, Marie Zufferey*, Daniele Tavernari*, Natalya Katanayeva, Marco Mina, Kyle M. Douglass, Timo Rey, Franck Reynaud, Suliana Manley, Giovanni Ciriello, and Elisa Oricchio

Published in Nature Genetics on May 29, 2019; volume 51, pages 517–528
DOI: 10.1038/s41588-018-0338-y
(Appendix 1)[44]

Using high-throughput chromatin conformation capture (Hi-C), we found that epigenetic changes induced by *EZH2*^{Y646X} do not disrupt hierarchical chromatin organization, but rather are enriched within chromatin topologically associating domains (TADs). In particular, by mapping transcriptional changes to chromatin organization in lymphoma, we found that the increase of H3K27me3 in *EZH2*^{Y646X} tumors drives concordant gene silencing within TADs, leading to the inactivation of entire domains (termed inactive TADs). Inactive TADs are enriched for *EZH2* target genes involved in b-cell differentiation and cell death. In one specific example, we show that synergistic inactivation of multiple tumor suppressors within the same inactive TAD, TAD6.139, synergistically accelerates b cell proliferation and lymphoma progression *in vivo*. Indeed, pharmacological inhibition of mutated *EZH2* depletes H3K27me3, restores intra-TAD chromatin interactions to the status observed in *EZH2*^{WT}, and reactivates transcriptional activity within TAD6.139. Finally, using targeted chromatin conformation capture with a unique molecular identifier (UMI-4C) for TAD6.139, we found that the intra-TAD epigenetic and transcriptional changes, driven by *EZH2*^{Y646X}, are coupled with rewired promoter-promoter interactions. Using STORM super resolution microscopy, we observed structural modifications of TAD6.139 associated with the re-wiring of promoter interactions of DNA. Our results underline the relevance of the genome topology in understanding how cancer-associated epigenetic alterations act beyond single genes and instead modify the activity of multiple genes comprised within entire chromatin domains.

2.4 Discussion

The PRC2 complex binds up to 20% of all gene promoters found in mammals, making it one of the master regulators of transcriptional silencing in the cell. PRC2 genes are often mutated in cancer, and correlate to disease progression, metastasis, and poor prognosis in patients[71, 90]. *EZH2* gain-of-function mutations (*EZH2*^{Y646X}) in GCB-DLBCL and follicular lymphoma were first characterized in 2010 and have since become an intense area of interest for targeted therapeutic approaches in the clinic[107, 114, 119, 125]. Indeed, the characterization of *EZH2*^{Y646X}, as well as other mutations identified in the SET domain of *EZH2*, lead to the development of a number of small molecule competitive inhibitors of S-adenosylmethionine (SAM), the methyl donor essential for the HMT activity of *EZH2*^{Y646X}, and have showed some promise in early phase clinical trials (NCT02875548; NCT01897571)[78]. However, some limitations with regard to efficacy and acquired resistance remain problematic (NCT02082977: terminated). Understanding the underlying mechanisms of *EZH2*^{Y646X} mutations in cancer development and disease progression will provide a more complete picture to inform the clinical efficacy of attacking this target in patients.

Our study takes into account the three-dimensional organization of chromatin and its fundamental building blocks, DNA and histones. The process of folding DNA inside the nucleus is not random, rather DNA is packaged systematically to position genes and their transcriptional machinery into the exact right location to turn transcription on or off with ease in a tightly regulated manner. Interestingly, this process is not static, but can be dynamic, meaning that chromatin can move and shift to adjust to what it needs at any given time while also maintaining a remarkable memory function that can be passed through generations of the cell cycle. While this process has been well studied under normal homeostatic conditions (during cell fate commitment, for instance), we do not fully understand how this delicate balance is perturbed when a normal cell becomes malignant.

In our study, we begin unraveling the complexities of this delicate balance between transcriptional programming and the three-dimensional organization of chromatin, and how this goes awry in the context of *EZH2*^{Y646X} mutations in lymphoma. By taking into account the three-dimensional organization, we were able to pinpoint a novel mechanism of epigenetically-mediated transcriptional hijacking that occurs in lymphoma, which is coupled with the structural rearrangement of the chromatin itself.

EZH2^{Y646X} mutations lead to a global increase of H3K27me3 and subsequent transcriptional silencing within the cell. By mapping the epigenetic and transcriptional profiles of *EZH2*^{WT} and *EZH2*^{Y646X} mutant lymphomas to the three-dimensional organization of chromatin, we characterized the targeting and synergistic inactivation of entire chromatin domains (termed inactive TADs) by *EZH2*^{Y646X}. We functionally validate target TAD6.139, which contains two tumor suppressor genes: *FOXO3* and *SESN1*. Interestingly, both *FOXO3* and *SESN1* are direct transcriptional targets of the tumor suppressor gene, p53[19, 145], and have a functional role in Akt-mTOR signaling. *FOXO3* is a transcription factor directly regulated by Akt[39], and *SESN1* regulates mTOR activity and its downstream translational program[134]. Indeed, we found that concurrent inactivation of *FOXO3* and *SESN1* synergistically promotes B-cell proliferation *in vitro* and tumor growth *in vivo*. By mapping the transcriptional changes associated with *EZH2*^{Y646X} epigenetic alterations to chromatin structure, we were able to elucidate the synergistic

behavior between these two genes, a phenomenon that would have likely been missed otherwise.

Importantly, we were able to rescue this effect using small molecule inhibitors (EPZ-6438 and GSK-126) targeting *EZH2 in vitro*. Using a specialized chromatin conformation capture technique to target interactions occurring within TAD6.139 (UMI-4C), we were able to pinpoint structural changes within the domain. By analyzing the chromatin interactions within TAD6.139 in GSK-126 treated and untreated *EZH2*^{Y646X} lymphoma cell lines, we found that upon TAD re-activation, the chromatin conformation is modified to bring the *FOXO3* and *SESN1* promoters into contact with one another, thus allowing transcription to proceed in a coordinated fashion. We further confirmed these structural changes by visual inspection using STORM super-resolution microscopy. We conclude that this structural modification is an important factor in the transcriptional regulation of these tumor suppressive genes, providing a functional link between histone post-translational modifications, chromatin structure, and transcription in the context of *EZH2*^{Y646X} lymphoma.

Our work uncovering this novel mechanism of action of *EZH2*^{Y646X} identifies a potential secondary effect of targeting *EZH2* in the clinic that was not previously known. For example, the genes found within TAD6.139 can be used as novel biomarkers for the efficacy of *EZH2*-targeted therapy in patients, or could be considered potential candidates for targeted therapeutics themselves. Interestingly, TAD6.139 is often deleted in lymphoma, and these deletions tend to be mutually exclusive with *EZH2*^{Y646X} mutations, leading us to speculate that the synergistic down-regulation or chromosomal deletion is sufficient to block this distinct tumor suppressive activity and is important in this context.

In summary, our study demonstrates that mutations in the epigenetic modifying enzyme, EZH2, can lead to modifications in the structure and function of entire topologically associating domains. In the context of lymphoma, *EZH2*^{Y646X} mutations target b-cell specific and tumor suppressive TADs to synergystically downregulate transcription and disable their tumor suppressive function.

2.5 Further considerations and future directions

2.5.1 Further studies of EZH2^{Y646X} in the context of b-cell lymphoma

In future studies, it will be interesting to further elucidate the functional role(s) of *EZH2*^{Y646X} in b-cell malignancies in the context of chromatin structure. In our study, we identify a subset of chromatin domains (n=72) that are epigenetically and transcriptionally down-regulated by *EZH2*^{Y646X}. In future studies, it will be interesting to determine the functional relevance of these other *EZH2*^{Y646X} target regions. In addition, while we see that the subset of inactive TADs identified in our study are enriched for various b-cell specific and tumor suppressive programs, their precise relevance is still yet to be determined. We provide a functional mechanism for the tumor suppressive TAD6.139, but it will be interesting to further interrogate other *EZH2*^{Y646X} target TADs in order to determine their functional relevance in lymphoma pathology. In addition, our analyses identified a small number of domains which actually become transcriptionally activated in *EZH2*^{Y646X} lymphoma (n=2). While much less abundant than their inactive counterparts, it would be interesting to determine if this activation is directly caused by *EZH2*^{Y646X}, or indirectly by some unknown mechanism, and if these TADs are functionally relevant.

2.5.2 What is the mechanism of action of EZH2^{Y646X} in other cancer subtypes?

EZH2^{Y646X} mutations are observed in other tumor subtypes, albeit at a much lower incidence than observed in b-cell lymphomas. In b-cell lymphoma, *EZH2*^{Y646X} target TADs are enriched for tissue specific programs, so it is intriguing to think that *EZH2*^{Y646X} may target different domains (tissue-specific domains) in tumors of other tissue types. If so, what would drive this conscience decision-making? It is intriguing to speculate that it stems from the cell-lineage specific properties of PRC2, and cell epigenetic memory may play a role. For example, it has already been shown that *EZH2*^{Y646X} is sufficient to drive tumorigenesis *in vivo*. If *EZH2*^{Y646X} oncogenic mutations are introduced to one of two cell lineages derived from the same stem cell population (lymphocytes versus melanocytes, the two most common tumor cell types with *EZH2*^{Y646X}, for instance), are the target regions of *EZH2*^{Y646X} tissue specific? And to what degree is the target specificity conserved within a given tissue? Finally, how would this contribute to tumorigenesis in a tissue-specific context?

2.5.3 What is the link between EZH2^{Y646X} and chromatin structure?

We provide a functional link between *EZH2*^{Y646X} and chromatin structure through TAD6.139, however the exact mechanism of PRC2-mediated chromatin looping is still unclear. Indeed, the global mechanisms guiding changes to chromatin architechture are still unclear in the context of PRC2. In our study, we see a clear link between H3K27me3 and chromatin structure, however we do not yet understand the molecular mechanisms underlying this apparent synergy. Which begs the question, to what extent can PRC2 modify chromatin structure, and what are the features of PRC2 that work to drive chromatin compaction in this context? It has been suggested that PRC2-driven chromatin looping is actually mediated by PRC1, however this mechanism has not been well established in mammalian cells, even under normal homeostatic conditions or during development. Previous studies of mouse ES differentiation provide an impressive link between PRC1 and chromatin compaction, but demonstrate that this mechanism is not dependent on H2AK119Ub[95]. Instead, one hypothesis is that there may

be a distinct PRC1 complex conformation (or subtype) that could mediate chromatin compaction. As previous studies have shown, there are 16 known PRC1 subunits, making at least 180 potential PRC1 conformations. Thus, there may be a distinct PRC1 associated protein or subunit conformation that lends the mechanical force to alter chromatin structure. In addition to PRC1, previous studies have shown the ability of the (BAF) SWI/SNF complex to remodel nucleosomes in an ATP-dependent manner. Indeed, in studies of mammalian cell differentiation, it was shown that PRC2 and BAF subunit proteins counterbalance one another during mouse ES-cell differentiation, indicating a functional interplay between the two complexes during development[222]. In addition, the BAF complex has been implicated in HSC maintenance, and regulation of erythroid, lymphoid, and myeloid lineages in studies of mouse ES cells. Furthermore, it has been shown that the BAF complex plays a role in chromatin remodeling during hematopoesis, and aberrant expression of BAF complex proteins have been previously implicated in hematologic malignancies[220]. In the context of PCR2-*EZH2*^{Y646X}-driven chromatin remodeling, I hypothesize that additional chromatin remodeling factors, such as PRC1 and BAF, among others, may have some secondary role in the structural remodeling of the *EZH2*^{Y646X}-target chromain regions, and further functional studies combining epigenetics and transcriptomics with analyses of chromatin 3D structure will further elucidate these interdependencies.

2.5.4 What is the functional relevance of other genomic alterations in the context of chromatin organization in lymphoma?

Multiple genomic alterations, including mutations, chromosomal rearrangements, and epigenetic alterations can contribute to cancer development and progression. Many alterations observed in cancer cover large portions of the genome encompassing not just one single gene, but multiple genes. When an alteration spans large portions of the genome it can be difficult to understand the functional relevance of such alterations. it has been shown that chromosomal alterations in cancer can modify the three-dimensional structure of chromatin to drive oncogenic programs[217]. Using the abundance of data generated from genome-wide transcriptional, epigenetic, and Hi-C experiments in Donaldson Collier et al., 2019, it is feasible to begin to explore the relevance of other genomic alterations in DLBC lymphoma on chromatin structure and pathogenic function.

In a follow-up study of chromatin dynamics in cancer, we have begun to explore the functional relevance of the chr.2p16.1 amplification in the context of b cell malignancies. Amplifications of Chr.2p16.1 are observed in 18-26% of b cell lymphomas, and always encompass two genes: *BCL11a* and *REL*. *BCL11a* is a transcription factor essential for lymphoid development[106], and *REL* is part of the NF-κB complex that regulates the gene expression in activated b-cells[210]. Regardless of the prevalence of this amplification, the functional link between chr.2p16.1 amplifications and lymphomagenesis remain illusive.

Using analyses of chromatin structure, we are investigating the effects of the chr.2p16.1 amplification on chromatin organization and chromatin interactions at different levels of hierarchical organization of the genome. Using gene expression analysis, we can map the differentially expressed genes to structurally altered genomic regions, allowing for the identification of genes that are both directly and indirectly affected by the chr.2p16.1 amplification. Finally, using chromatin-immunoprecipitation combined

with high-throughput sequencing (ChIP-seq), we can map the exact location of enhancers (marked by H3K27ac) and promoters (marked by H3K4me3) in b cell lymphomas with or without the chr.2p16.1 amplification. In future studies, we will further elucidate how changes in chromatin interactions affect gene expression in cancer. This particular follow up study will integrate chromatin structural changes with oncogene expression, thus providing a new lens to understand the role of chromatin three-dimensional organization in supporting tumor growth.

Chapter 3: Dissecting epigenetic, transcriptomic, and metabolic signaling dependencies using a novel model of central nervous system primitive neural ectodermal tumors (CNS-PNETs)

3.1 Introduction

3.1.1 Pediatric Central Nervous System Tumors

Pediatric central nervous system (CNS) tumors are the leading cause of cancer-associated death in children worldwide, notwithstanding the lower frequency compared to other tumor subtypes, including leukemia and lymphoma (WHO). Despite many hallmark advances to our understanding of the mechanisms of development, growth, and treatment of cancer over the last 50 years, advances in our understanding of pediatric CNS tumors has lagged in comparison.

Pediatric CNS tumors, on average, harbor a 14-fold lower mutational load than adult tumors of a similar tissue of origin[206]. For many of these tumors, where driving mutations cannot be identified, it is unclear how and under which conditions these tumors arise. Furthermore, relatively few models exist to explore the underlying mechanisms of tumor development, progression, and response to treatment[123]. It has been proposed that pediatric CNS tumors arise as a result of aberrant neuronal development, with both the cell of origin and anatomical location contributing to tumor behavior[169]. There is evidence that many CNS tumors propagate not necessarily as a result of genetic mutations, but rather from epigenetic alterations - particularly aberrant DNA methylation patterning[208]. In addition, pathological and molecular analyses of pediatric CNS tumors consistently reveal high expression of various neural stem cell and neural development markers, including the neural stem cell marker, CD133 and the neural development type VI intermediate filament protein, Nestin, indicating that tumors may arise from a small population of neural cancer stem cells that have the ability to generate and propagate tumor development[169,188]. Cancer stem cells are quite similar to normal stem cells in that they posses the typical characteristics of self renewal and the ability to differentiate into multiple cell types, however unlike normal stem cells, their capacity for self-renewal is uncontrolled. The data suggests it is likely that both mechanisms play an important role in pediatric CNS-tumor initiation and development, as well as resistance to treatment. However, additional research is required to better understand these mechanisms of action across various subtypes of pediatric CNS tumors.

3.1.2 Pediatric Central Nervous System-Primitive Neural Ectodermal Tumors (CNS-PNETs)

Central nervous system primitive neural ectodermal tumors (CNS-PNETs) are a rare but rather aggressive subtype of pediatric CNS tumors arising from poorly differentiated neuro-epithelial cells. Due to the rarity of this tumor, comprising of approximately 5% of all pediatric brain tumors, significant advances in new diagnostic tools, molecular-based characterization, and therapeutic strategies have been limited

in scope[123]. While CNS-PNETs histologically resemble medulloblastomas and high grade gliomas, previous studies have shown that these tumors have distinct molecular profiles that are representative of tumors arising from different cells of origin[79, 103, 138]. In addition, the cytogenic profiles and genetic alterations observed in CNS-PNETs are remarkably different from those observed in other CNS tumors, indicating unique biological derivations between tumor subtypes that are not otherwise apparent on the basis of morphological features alone[103]. However, most of these studies are descriptive and the paucity of specific *in vitro* and *in vivo* CNS-PNET models has limited the possibility to define the role of these genomic alterations in tumor development and progression.

3.1.3 Diagnosis and Treatment of CNS-PNETs

Traditional diagnosis of pediatric CNS-PNETs is rather difficult and sometimes even inaccurate, relying solely on anatomical location and descriptive histological features [23, 79, 208] . Based on the 2007 WHO report of pediatric CNS tumors, CNS-PNETs are defined as highly heterogeneous tumors that exhibit remarkably divergent patterns of differentiation. CNS-PNETs generally encompass tumors containing poorly differentiated cells that follow along neuronal, astrocytic, or ependymal lines of origin, which underlines the heterogeneity observed[108].

Although it is noted that CNS-PNET tumors are highly heterogeneous across patients and exhibit variable responses to treatment, regimens remain more or less consistent in practice, with little understanding of the underlying causes for such variability in prognostic outcome[169]. Treatment of CNS-PNETs is limited to intensified chemo and radiotherapy following surgical resection, a regimen initially designed for high grade medulloblastomas[21]. More recently, to reduce the neurological damage in young patients (<4 years of age), radiotherapy has been replaced with an intensive chemotherapy treatment regimen of five agents, including methotextrate, clophosphamide, vincristine, carboplatin, and etoposide (SKK-HIT'92 clinical trial)[148]. However, CNS-PNETs and medulloblastoma are molecularly distinct from one another, and from a prognostic standpoint, CNS-PNETs show a significantly worse clinical outcome, with an estimated 3-year progression free survival in only ~50% of patients regardless of the intensified treatment regimen[79, 170]. Surgical resection, when possible, remains the best prognostic measure for progression-free survival and overall survival[79]. However, complete macroscopic surgical resection is only feasible in 30-50% of patients, oftentimes due to the tumor's precarious location[79, 170]. Furthermore, despite their small numbers, these highly aggressive and highly resistant tumors primarily affect young children, whereby if they survive the disease they may suffer long-term damage due to the chemo- and radio-therapeutic treatment[88].

In the context of CNS-PNETs, few advances have been made to apply targeted novel therapeutic approaches to the treatment of the disease, which may be in large part due to both the rarity of the disease and a general lack of research tools available to dissect the molecular mechanisms driving tumor initiation, development, and treatment response[123]. More recently, some clinical trials have been initiated to combine conventional chemotherapy treatment with targeted therapeutic strategies, such as Bevacizuman (Avastin) targeting vascular endothelial growth factor (VEGF) to inhibit angiogenesis (ClinicalTrial ID: NCT01217437). Nonetheless, given the unique complexity underlying CNS-PNET pathogenesis and the high mortality rates associated with this disease, there is a pressing need to improve

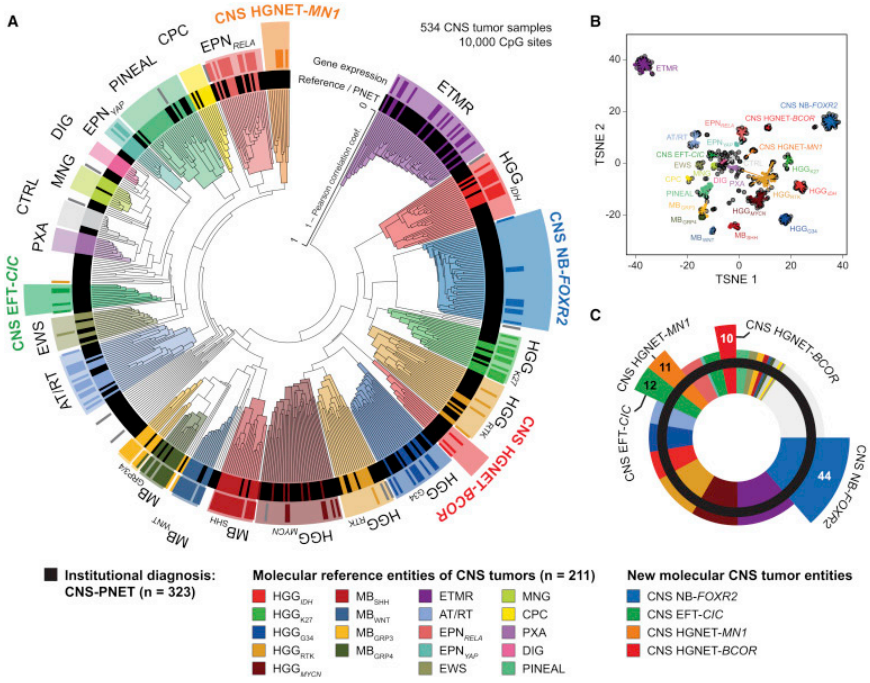
our knowledge to identify new diagnostic and therapeutic strategies in the clinic[123].

3.1.4 The Role of DNA Methylation in CNS-PNET Pathogenesis

DNA methylation is an epigenetic program whereby methyl groups are directly added to the DNA sequence to alter the transcriptional program of the cell without directly modifying the genetic code itself. Cancer cells often exhibit abnormal DNA methylation patterns compared to normal cells. In normal cells, individual CpG sites are often methylated, and CpG-islands (CGIs) are often unmethylatlated. In cancer, individual CpG sites become globally unmethylated (hypomethylation) with the exception of CGIs, which actually become hypermethylated[50]. It is thought that DNA hypomethylation is an early event in cancer, introducing chromatin instability and activating proto-oncogenes[47,40]. Cancer-driven hypermethylation of CpG islands (CGIs) is often located at promoters of tumor suppressor genes, which are frequently transcriptionally down-regulated as a result of their methylation status. The identification of this epigenetic regulatory network in cancer has opened a therapeutic window for prognosis and treatment applications in many cancer subtypes. Indeed, DNA methyltransferase (DNMT) inhibitors were some of the earliest epi-therapeutics approved for the treatment of cancer, and are currently used for the treatment of various hematologic malignancies[56, 57]. In addition, aberrant DNA methylation has been further linked to important signaling events in in various cancer subtypes. For example, estrogen receptor-α (ERα) activity was found to be tightly coupled with DNA methylation in breast cancer, where the methylation status of CpGs is correlated with the transcriptional activity of the hormone receptor and were found to be good predictors of clinical response to anti-estrogens[73, 172]. In summary, DNA methylation is an important mechanism for epigenetic and transcription regulation in the cell. Aberrant DNA methylation patterning is observed across many types of cancers, which functions to alter transcription and enact oncogenic signaling to drive tumor development and progression.

In the absence of defining genetic lesions present in CNS-PNETs, it has been hypothesized that these tumors may instead be driven by epigenetic alterations, specifically maladaptive alterations to the DNA methylation profiles that occur during development[206, 23, 208]. To this end, there has been a recent effort to profile the epigenetic landscape of pediatric CNS-PNETs, providing new insight into the unique biology underlying the development of these tumors. In a multi-institutional collaboration, researchers collected over 350 institutionally diagnosed pediatric CNS tumors biopsies for matched DNA methylation and expression analysis, thus creating the largest information repository of these rare tumors available to date[23,208]. The implementation of high-throughput sequencing methods has provided a unique opportunity for identifying epigenetic biomarkers even within rare, mutation-low tumor types like pediatric CNS-PNETs. The results from this study revealed remarkable heterogeneity between patient samples, with some CNS-PNETs exhibiting features of other types of CNS-tumors, as well as the identification of new unique sub-classes (Figure 3). This novel classification system is based specifically on the DNA methylation profile of each patient, indicating that epigenetic alterations may be a powerful driving force behind CNS-PNET development[208]. In light of this new finding, it will need to be determined whether this new classification system will lead to improvements in the diagnosis and treatment of patients in the clinic. Most importantly, the identification of specific biomarkers and prognostic factors relating to each distinct subtype will be important to help translate such findings to the clinical setting.

Figure 3: Molecular Classification of CNS-PNETs by DNA Methylation Profiling

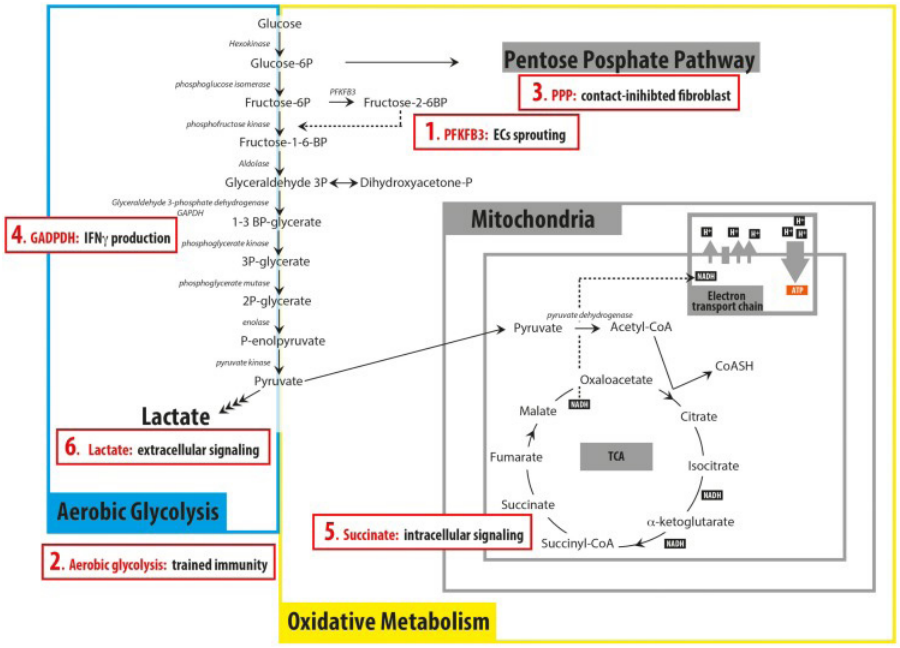


From: New Brain Tumor Entities Emerge from Molecular Classification of CNS-PNETs
Sturm et al., Cell 2016
DOI: 10.1016/j.cell.2016.01.015

3.1.5 Cancer Metabolism and Metabolic Imaging

Metabolism encompasses a set of biochemical processes that function to turn organic material into energy, synthesize new materials in the body, and eliminate waste; a process that is vital to every known living organism. Metabolic transformation is a major hallmark of cancer, whereby cells undergo a divergent switch to sustain uncontrolled growth and cell survival[64, 65]. The most common metabolic transformation was first observed by Otto Warburg in the 1920s, and is referred to as aerobic glycolysis or the "Warburg effect" [214, 215, 216]. In humans, differentiated cells generally undergo glycolysis to convert glucose into pyruvate to feed into the tricarboxylic acid (TCA) cycle in the mitochondria to produce adenosine triphosphate (ATP) via oxidative phosphorylation (Figure 4). However, many tumor cells instead convert glucose into lactate for the production of ATP even in the presence of oxygen (Figure 4), whereas in quiescent cells this process is usually only observed under anaerobic conditions. While, in the presence of oxygen, aerobic glycolysis is generally considered less efficient than mitochondrial-driven oxidative phosphorylation, cancer cells are able to produce ATP more rapidly to support continuous cell division, growth, and invasion[64, 65].

Figure 4: Schematic representation of the principal pathways involved in glucose catabolism



From: Aerobic Glycolysis: Beyond Proliferation
William Jones and Katiuscia Bianchi
DOI: 10.3389/fimmu.2015.00227

While metabolic transformation to aerobic glycolysis has been observed across many different tumor types, recent evidence has suggested that a number of additional metabolic processes can be altered in cancer as well. For example, recent studies have shown that certain cancers can also up-regulate the oxidative phosphorylation (OXPHOS) pathway, thus providing some rationale for treating tumors with OXPHOS inhibitors in specific contexts[4]. Other evidence has identified impaired mitochondrial function linked to specific mutations in mitochondrial proteins, or simply due to an increase in mitochondrial DNA content in some cancers[160].

One particularly intriguing point is that cellular metabolism can be highly dynamic and heterogeneous across tumor types or even within the same tumor. For example, one study of pancreatic ductal adenocarcinoma (PDAC) suggests that a population of cancer stem cells within PDAC tumors are more reliant on oxidative phosphorylation, while the surrounding non-stem cancer cells rely more heavily on

aerobic glycolysis[160, 174]. These results suggest that intra-tumor molecular heterogeneity is associated to distinct metabolic signatures. In another example, researchers used ^{1}H magnetic resonance spectroscopy (MRS) to profile the metabolic features of three subtypes of brain tumors reflecting various stages of aggressiveness and invasiveness. In each subtype, the metabolic signature contained unique features, with low-grade astrocytomas exhibiting high levels of myo-inositol, and more mildly elevated levels of lactate and choline. Conversely, glioblastomas exhibited the opposite, with high levels of lactate and choline and very low (or no) signal detected for myo-inositol[25]. These results indicate that tumors may undergo either: (1) dynamic switching that occurs during the evolution of tumor growth and transformation, or (2) indicate differences in metabolic signaling dependencies between tumor subtypes and/or cell of origin (for example: astrocytes and glial cells). In summary, metabolic transformation is an important hallmark of cancer, but the types of metabolic alterations can vary between tumors and even within tumors.

The abundance and range of metabolic alterations observed in cancer pinpoints an attractive avenue for treatment targeting cancer-associated metabolic signaling, as well as exploiting these fundamental differences for cancer detection and patient stratification. One of the most successful examples is the use of radioactive glucose derivatives combined with positron emission tomography (PET) imaging for detection of various types of cancer[22]. ^{18}fluorodeoxyglucose (FDG) is a glucose analogue, whereby upon uptake, FDG becomes phosphorylated and then trapped, leading to an intense radiolabeling of cells and tissues with high glucose uptake. FDG-PET is successful and routinely used in the diagnosis, staging, and monitoring treatment response in many types of cancer, including Hodgkin and non-Hodgkin lymphomas, colorectal cancer, lung cancer, and melanoma[12]. In the context of the PI3K pathway and its role in mediating glucose transport and metabolism, FDG-PET has even been used to measure the efficacy of treatment with PI3K-inhibition in xenograft mouse models of human cancer[48, 106, 113]. However, in tissues with high glucose uptake under homeostasis, like the brain, liver, and kidneys, FDG-PET measurements are not always reliable[126, 146].

Current protocols for treating pediatric brain tumors take into account a number of different risk factors including age at the time of diagnosis, radiographic imaging features, and histological features, all of which contribute to diagnosis. In the clinic, diagnosis of CNS-PNETs is limited to magnetic resonance imaging (MRI). However, morphological features obtained from an MRI are not always reliable in predicting prognosis and treatment response[208]. While FDG-PET has been successfully applied in the clinic for the detection of various tumor types, providing valuable information about tumor size, aggressiveness, and treatment response, FDG-PET is not feasible for diagnosing CNS-PNETs due to the high background of glucose in the brain. For this reason, there has been a recent push to apply ^{1}H MRS-based metabolomics to human brain tumors to identify and stratify tumors based on their metabolic profiles[3, 87, 121].

Anywhere from 12-18 different metabolites can be reliably measured with ^{1}H MRS depending on the measurement techniques used, and reports indicate that there are a number of altered metabolites in both adult and pediatric brain tumors that can be reliably measured using ^{1}H MRS[3, 25, 66, 73, 121, 139]. Based on previous studies of primary tumors, some specific metabolites have emerged as potential biomarkers for

diagnosing and monitoring CNS-PNETs in the clinic[28, 54, 55, 117], and herein some examples are outlined.

Some metabolic features are robustly observed across almost all brain tumors, including a decrease in N-acetyl aspartate (NAA) and an increase in choline[73, 179]. N-acetyl aspartate is one of the most abundant amino acids in the central nervous system and is thought to be a contribution from neurons, axons and dendrites, and loss of NAA is indicative of loss of normal neuronal density[7]. Choline is a major source of the methyl groups used in the synthesis of phosphatidylcholine and sphingomyelin, two important phospholipids for the generation of the cell membrane. Thus, choline is thought to be a biomarker for membrane synthesis and/or degradation, and elevated levels of choline could indicate increased membrane turnover and high cellular density[7, 63].

Other metabolic changes that have been observed in brain tumors include lactate, glycine, and myo-inositol, albeit more variable depending on the tumor subtype, size, and aggressiveness[25, 66, 73]. An increased level of lactate is generally thought to be a result of anaerobic glycolysis in tumor cells, however other factors may contribute to lactate abundance, including lack of blood flow in the region of interest as well as tissue necrosis[73]. Glycine is an important amino acid for nucleotide biosynthesis, particularly pyrimidine biosynthesis. Glycine/Serine metabolic pathway associated enzymes (glycine-N-methyltransferase (GNMT) and S-adenosylmethionine (SAMe)) have been linked to aberrant DNA methylation signatures in hepatocellular carcinoma, providing an interesting link between altered glycine metabolism and DNA methylation in cancer[76, 179]. In addition, glycine has been identified as a marker for poor prognosis in pediatric brain tumors[6]. Myo-inositol is primarily produced in astrocytes. A derivative of glucose, myo-inositol and is considered a secondary messenger in many signal transduction pathways and is considered an important osmolarity regulator in the brain. Relative to normal brain tissue, increases in myo-inositol have been linked to lower-grade brain tumors, like low grade gliomas and astrocytomas, but is decreased in high grade gliomas and glioblastoma[164], therefore making it an attractive marker associated with tumor aggressiveness. However, it is not clear if myo-inositol is linked directly to aggressive behavioral traits (e.g. increased in cellular proliferation and tissue invasion). Indeed, it may also be linked to distinct features of the tumor itself, which depend on factors suchs as, the cell or origin (astrocytes versus glial cells), or osmoregulation related to tumor-induced brain swelling and blood brain barrier damage.

In addition to ^{1}H MRS-based methods, live imaging using hyperpolarized magnetic resonance (HMR) has also gained a considerable amount of attention in recent years. HMR allows for the noninvasive measurement of metabolic reactions *in vivo*, following the administration of a ^{13}C-tagged substrate. Hyperpolarized magnetic resonance is >10,000 times more sensitive than previously used methods, such as ^{1}H MRS, which only has the ability to measure highly abundant amino acids[211]. MRS technology implies numerous metabolic intermediates between the injected precursor and many assumptions are required to derive the underlying enzymatic fluxes. Through the use of hyperpolarized magnetic resonance, the ^{13}C signal is enhanced by several orders of magnitude, which allows one to measure metabolic processes with much higher spatial and temporal resolution. As a result, HMR gives access to the intermediates present in various metabolic pathways, and determines the enzymatic flux of each reaction that are otherwise undetectable[33].

In conclusion, given the paucity of methods for identifying cancer risk factors combined with the established patterns of metabolic alterations observed in brain tumors, non-invasive, *in vivo*-based metabolic imaging techniques lends an attractive method for stratifying behavioral characteristics for the refinement of risk factors associated to the diagnosis and treatment of CNS-PNETs.

3.1.6 Existing Models of CNS-PNET Pathogenesis

Due to the rarity of these tumors and the lack of knowledge of oncogenic drivers, few *in vitro* and *in vivo* models exist for CNS-PNETs. To our knowledge, one patient-derived CNS-PNET cell line exists, called PFSK-1 (ATCC CRL-2060)[53]. Given the heterogeneity observed in primary CNS-PNETs, the characteristics of one cell line cannot encompass the global behavior of CNS-PNET tumors, and thus the development of further models is required to fully capture the biology of these tumors. More recently, an *in vivo* zebrafish model of CNS-PNETs was developed to encompass the broad heterogeneity of CNS-PNET tumors[118]. However, this model is based on non-human cell of origin, and may not accurately recapitulate the distinct human biology of CNS-PNETs. Indeed, only 70% of human genes share orthologs with zebrafish genes, and thus there may be clear divergences in zebrafish models of human genetics that do not accurately translate to the genetics of human diseases[18]. In addition, while a small number of primary patient-derived xenograft (PDX) mouse models do exist for CNS-PNETs, these models have not been publically available, thus limiting their potential to be used for pre-clinical research of CNS-PNETs globally.

3.2 Rationale for Study

It has been speculated that CNS-PNETs arise from poorly differentiated or undifferentiated neuro-epithelial cells. In the absence of distinct genomic alterations, it is thought that these tumors are instead driven by aberrant DNA methylation patterning that works to favor oncogenic programming in the cell, which gives rise to malignant transformation. Furthermore, treatment options are very limited and prognosis is poor for patients suffering from this disease, pinpointing a pressing need to find alternative methods for diagnosing, stratifying, and treating patients in the clinic. Given that CNS-PNET tumors arise from poorly differentiated neural epithelial cells that resemble the early stages of neural differentiation and the paucity of specific *in vivo* models for studying CNS-PNET pathogenesis, we wondered if it would be possible to derive a CNS-PNET mouse model based on human neural progenitor cells. We sought to develop a murine model based on neural progenitor cells derived from human induced pluripotent stem cells in an attempt to unravel disease complexities using a multi-faceted approach involving epigenetic, transcriptomic, and metabolic analyses.

3.3 Study Aims

1. To generate a robust model of CNS-PNETs based on hIPS-derived neural progenitor cells (human CNS-NPC model)

2. Identify the epigenetic, molecular, and physical features of CNS-NPC model to determine its accuracy in recapitulating key features of primary CNS-PNETs

3. To utilize the CNS-NPC model for pilot studies in the efficacy of metabolic imaging platforms to assess the efficacy of metabolic imaging platforms in the diagnosis of CNS-PNETs

3.4 Results: Modeling CNS-PNETs in vivo

"A model is a lie that helps you see the truth."
Siddhartha Mukherjee, The Emperor of All Maladies

3.4.1 Modeling CNS-PNETs in vivo

The lack of specific *in vitro* and *in vivo* CNS-PNET models has limited the opportunity to investigate the underlying mechanisms driving CNS-PNET tumor formation in order to identify novel ways of targeting these tumors using novel therapeutic intervention stragegies. Given that CNS-PNET tumors arise from poorly differentiated neural epithelial cells that resemble the early stages of neural differentiation, we wondered if it would be possible to derive a CNS-PNET mouse model based on human neural progenitor cells (NPCs). Previous studies have shown that it is indeed possible to recapitulate features of CNS-PNETs from neural progenitor and stem cells, however these models are based on NPCs derived from zebra fish, rather than human cells, and may lack some critical information regarding inherent biological differences between the two organisms[118]. To this end, we generated a xenograft CNS-PNET mouse model based on NPCs derived from human-induced pluripotent stem cells (hIPS cells).

3.4.2 Development of a CNS-PNET Mouse Model Derived from Human Stem Cells

hIPS cells were generated from IMR-90 fetal lung fibroblasts (ATCC CCL-186) by viral transduction of four genes: *OCT4, SOX2, MYC* and *KLF4*[129]. Following de-differentation into hIPS, cells are then re-differentiated towards a neural stem cell fate by a modified dual SMAD-inhibition protocol as previously described[168, 93], where in approximately 3-4 weeks neural progenitor cells are generated *in vitro* (Figure 5A and 5B). Using immunofluorescence and quantitative-PCR (qPCR) analyses of specific neural development markers, we were able to confirm the robust generation of NPCs *in vitro*. Indeed, NPCs *in vitro* show an up-regulation of neural progenitor markers, including *NESTIN*, *GFAP*, *PAX6*, *SOX2*, and Synaptophysin (*SYP*), at the mRNA and protein level (Supplemental Figure 5A and 5B). Upon differentiation, NPCs transduced with a plasmid containing green fluorescent protein (GFP) and luciferase are dissociated, collected, and intracranially transplanted into immune-compromised NOD-SCID mice (Figure 5C). The constitutive expression of GFP allows for cell identification in downstream histological analyses, and the luciferase allows for bioluminescence imaging (BLI) by IVIS spectrum (perkin elmer) by exploiting the light emitted following the chemical reaction between luciferase and its substrate, luciferin (Figure 5C and 5D)[201].

3.4.3 Analysis of the PFSK-1 Patient-Derived Cell Line

To our knowledge, one immortalized patient-derived CNS-PNET cell line exists, called PFSK-1 (ATCC CRL-2060)[53]. While the activity of one cell line does not necessarily represent the majority of the behavioral characteristics of all CNS-PNET tumors, it can be informative in some *in vitro* and *in vivo* experimental studies of CNS-PNETs and is used in this study as a second model of CNS-PNET pathogenesis. Using immunofluorescence and quantitative-PCR (qPCR) analyses, we characterized the expression of neural development markers in PFSK-1 cells *in vitro* (Supplemental figure 5A and 5D). mRNA expression and immunofluorescence protein expression analysis reveals that PFSK-1 cells exhibit lower expression of *NESTIN*, *SYP*, *GFAP, PAX6*, and *SOX2* relative to NPCs *in vitro*, indicating

that these cells may represent a different stage of neural differentiation. Nonetheless, given that these cells are derived from a primary CNS-PNET, we continued to use PFSK-1 cells as a second model of CNS-PNET pathogenesis.

3.4.4 Growth Characteristics of CNS-NPCs and PFSK-1 in vivo

BLI imaging of NPCs and PFSK-1 cells *in vivo* reveals that the growth tendencies between the two models is highly varied (Figure 5E). CNS-NPCs show some cell death within the first two weeks of implantation, followed by consistent growth over a period of approximately 8-10 weeks (Figure 5D). Conversely, PFSK-1 cells grow more rapidly, exhibiting highly aggressive behavior, with animals showing significant weight loss and behavioral abnormalities within three weeks of injection of only 10,000 cells (Figure 5D). Once tumors reached a size equivalent to or larger than $2x10^4 p/s/cm^2$ (normalized to the start of the experiment), as measured by IVIS imaging, tissues were collected for downstream analysis of DNA methylation, mRNA expression, and histological profiles to assess the similarities of the CNS-NPC model and PFSK-1 cells with primary patient samples.

3.4.5 Pathological Feature Analysis of CNS-NPCs and PFSK-1 in vivo

First, we assessed the molecular and physical characteristics relevant to CNS-PNET pathology through the histological analysis of CNS-NPC and PFSK-1 tumors. CNS-NPCs exhibit many of the typical physical characteristics of CNS-PNETs, including poorly differentiated small round cells, high nuclear-to-cytoplasmic ratio, with some Homer-Wright rosettes present (Figure 5 Fi). In addition, CNS-NPCs robustly exhibit high expression of the neural development type VI intermediate filament protein (Nestin) and the glial fibrillary acidic protein (GFAP), which are used as diagnostic markers of CNS-PNETs (Figure 5iii & 5iv, Figure 5G, Supplemental figure 5E). The proliferation index, as measured by Ki-67, is moderate, with approximately 40-60% of cells staining positive (Figure 5 Fii, figure 5G, supplemental figure 5E). Conversely, tumors originating from PFSK-1 cells exhibit a more uniform morphology, with a high nucleus-to-cytoplasm ratio, no Homer-Wright rosettes present, and large areas of vascularity as well as necrosis (Figure 5 Fi). In addition, PFSK-1 tumors exhibit moderate expression of Nestin (Figure 5 Fiii), are negative for GFAP (Figure 5 Fiv, figure 5G, supplemental figure 5E), and have a high proliferation index with 80-100% of cells showing active proliferation as measured by Ki-67 (Figure 5 Fii and figure 5G, supplemental figure 5E).

Next, we wanted to assess the impact of the neural microenvironment on the ability of NPCs to transform into CNS-NPC tumors. To address this point, we injected NPCs (Flank-NPCs) and PFSK-1 cells subcutaneously into the flank of immune-deficient NOD-SCID mice. Growth of flank-NPCs and flank-PFSK-1 was monitored by BLI-IVIS over a period of seven weeks. Growth of NPC and PFSK-1 cells were both impaired when injected into the flank, even considering the photon emission differences inherent to imaging through the brain and skull versus directly underneath the skin of the mice (Figure 5I). While PFSK-1 retained the same pathological features observed in the brain, flank-NPCs exhibited markedly different molecular and physical characteristics. Flank-PFSK-1 retained their dense nuclear-to-cytoplasmic ratio, high proliferation rate, and Nestin positivity (Figure 5 Hi, Hii, and Hiii). However, flank-NPCs lost many of these features, including a reduction in Nestin expression and Ki-67 (Figure 5 Hii and 5 Hiii), as well as some of the distinct physical characteristics observed in their

cerebral counterparts, with more diffuse nuclear content and higher percentage of infiltrating stroma (Figure 5 Hi). Overall, the NPC flank tumors exhibited loss of neural development-specific markers and a reduction in Ki-67, as well as differences in morphology compared to CNS-NPCs. Thus, when NPCs are injected into an alternative location in the body, these cells no longer retain the same features observed in CNS-NPC tumors, indicating tissue-specific priming is required for tumor development in this model.

In assessing morphological and physical characteristics, we identified both similar and distinct features of the CNS-NPC and PFSK-1 lesions that are associated with CNS-PNET pathogenesis. In the context of the CNS-NPC lesions, there is some brain specific priming that is required for development of CNS-PNET tumors. We conclude that both PFSK-1 and CNS-NPC models have distinct and relevant pathological features which can be exploited for further analyses of CNS-PNET tumor biology.

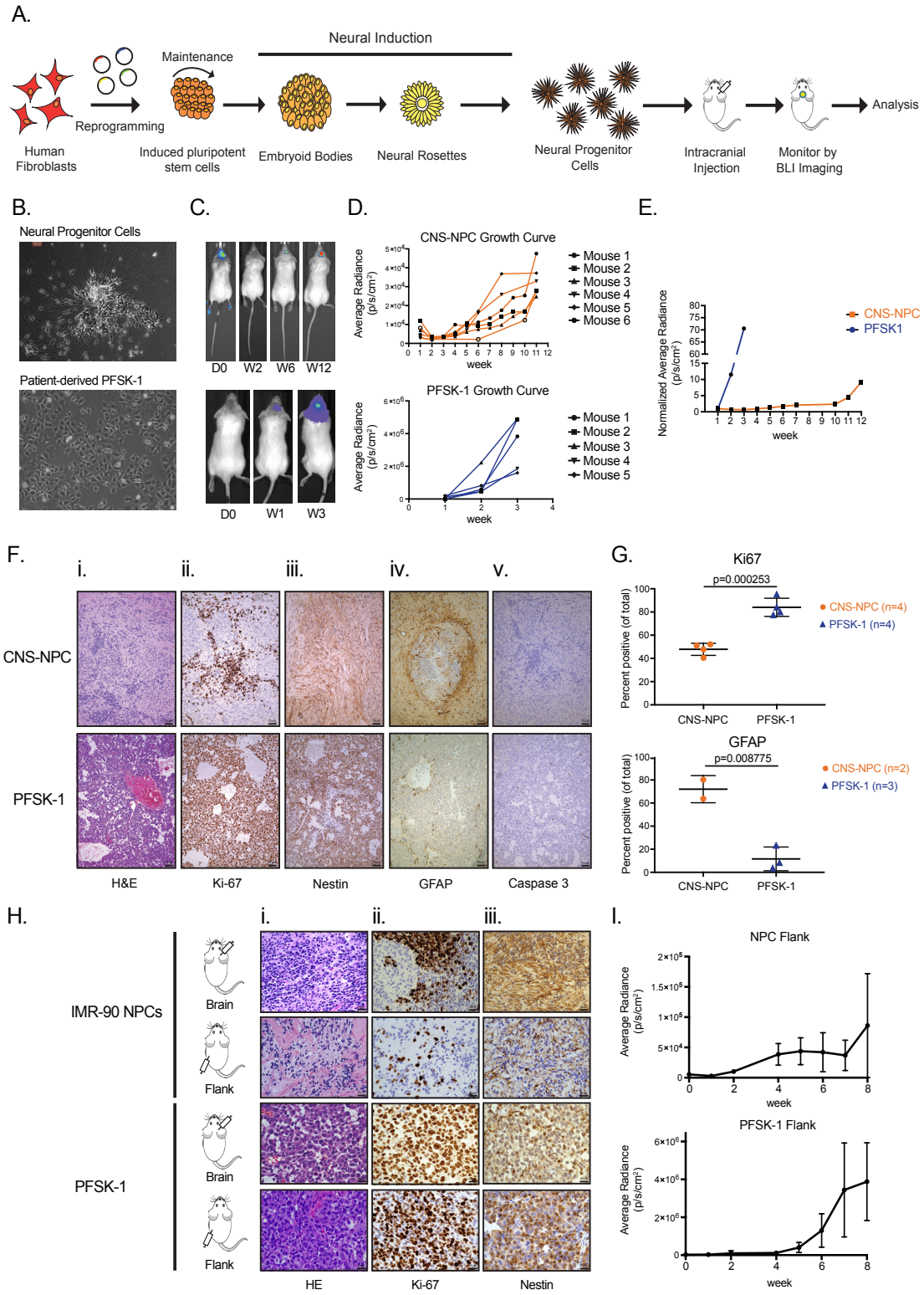
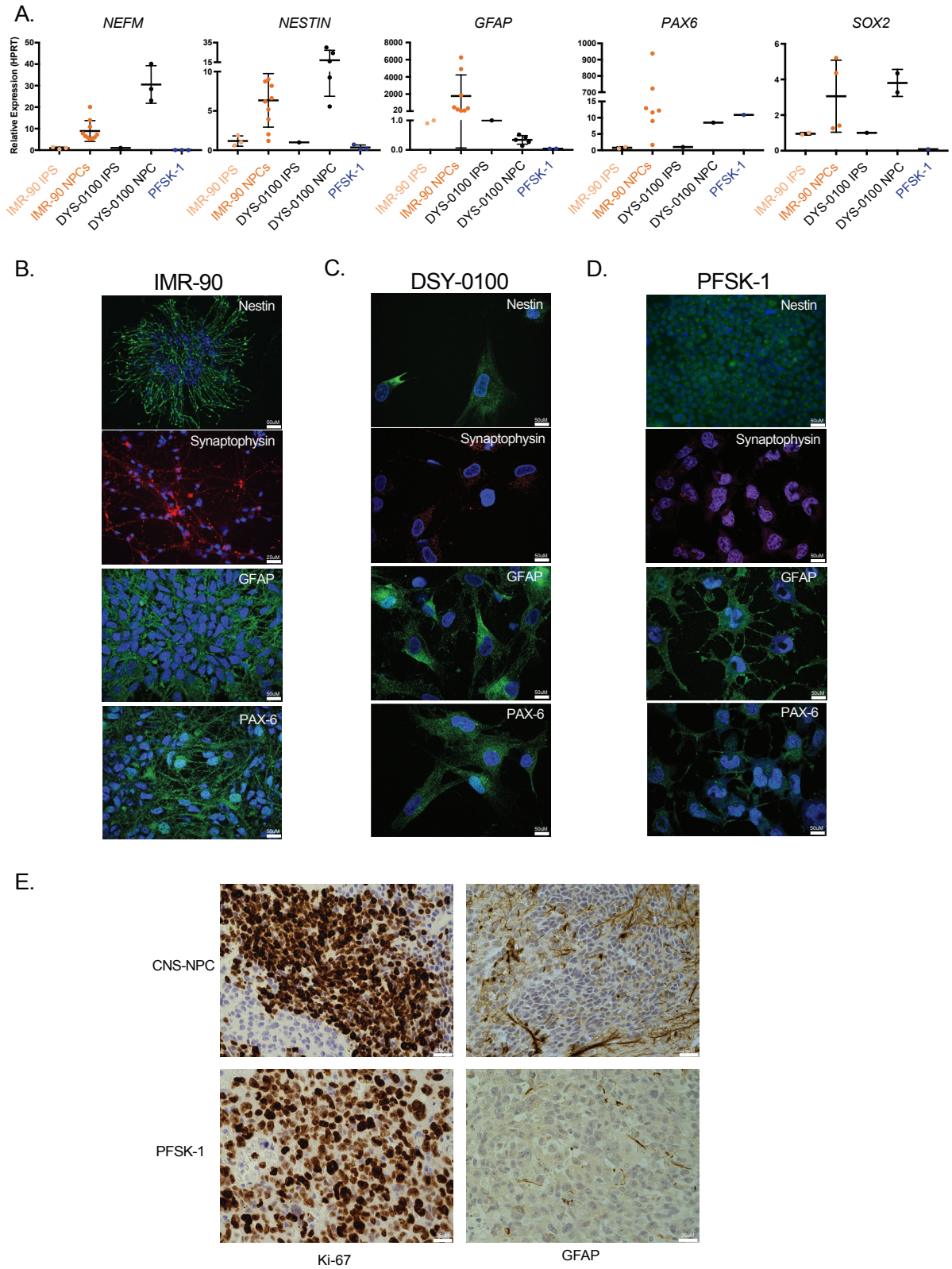


Figure 5. An in vivo Model for Studying CNS-PNET Tumors

(A) Schematic of the generation CNS-NPC tumors (B) Bright field images of NPC and PFSK-1 cells in vitro; 4X mag. (C) Representative BLI imaging of CNS-NPC and PFSK1 tumor bearing NOD-SCID mice over time. (D) CNS-NPC and PFSK-1 growth over time measured as the average BLI radiance (p/s/cm^2). (E) Representative growth curve of PFSK-1 and CNS-NPC tumors in vivo, measured as the average BLI radiance (p/s/cm^2), normalized to day 1 of injection. (Fi-v.) Reprensentative bright field images of immunostaining for (i) Hematoxylin & Eosin (H&E), (ii) Ki-67, (iii) Nestin, (iv) GFAP, and (v) Caspase-3 in CNS-NPC and PFSK-1 tumors; 10X mag. (G) Quantification of percent (%) Ki-67 and GFAP positive cells observed in CNS-NPC or PFSK-1 tumors. (Hi-iii) Representative bright field images of immunostain for (i) Hematoxylin & Eosin (H&E), (ii) Ki-67, (iii) Nestin in NPC and PFSK-1 cells injected subcutaneously into NOD-SCID mice; 40X mag. (I) The average BLI radiance (p/s/cm^2) observed in NOD-SCID mice subcutaneously injected into the flank with NPC or PFSK-1 cells.



Supplemental Figure 5. An in vivo Model for Studying CNS-PNET Tumors

(A) Average mRNA expression of *NEFM, NESTIN, GFAP, PAX6, SOX2* in IMR-90 hIPS, IMR-90 NPC, DSY-0100 hIPS, DSY-0100 NPC, and PFSK-1 cells in vitro. (B,C,D) Representative immunofluorescence images of (B) IMR-90-derived NPCs (C) DSY-0100-derived NPCs (D) PFSK-1 cells in vitro stained for Nestin (Green), Synaptophysin (SYP) (Red), GFAP (Green), and PAX6 (Green), and DAPI (Blue). (E) Representative image quantification counts for CNS-NPC and PFSK-1 tumors immunostained for Ki-67 and GFAP.

3.5 Results: Epigenetic and Molecular Analysis of CNS-PNETs

3.5.1 Epigenetic Analysis of CNS-PNETs

To analyze the DNA methylation and RNA expression profiles in CNS-NPCs and PFSK-1 cells, we isolated lesions from the surrounding tissue by enzymatic digestion and negative depletion of human cells. Isolated human cells were then processed for isolation of either genomic DNA or RNA. To exclude the possibility of contaminating mouse cells, we analyzed the RNA isolated from each lesion using mouse or human specific primers and RT-PCR and confirmed efficient selection of human cells, with negligible or no contamination from mouse cells (Supplemental figure 6A).

We performed bisulfite conversion on genomic DNA isolated from CNS-NPC or PFSK-1 cells and performed DNA methylation profiling using the HumanMethylation 850k BeadChip methylation microarray (Infinium). We performed a systematic comparison of the DNA methylation profiles of CNS-NPCs and PFSK-1 cells *in vivo* with the DNA methylation profiles of 323 primary CNS-PNET samples and 8 normal human brain reference samples[208]. First, principal component analysis of PFSK-1 cells and hIPS, NPCs *in vitro*, and CNS-NPCs shows the two models have markedly different methylation profiles (Supplemental figure 6B). Comparing the most differentially methylated probes (n=15,000) in primary patient samples, unsupervised clustering revealed that CNS-NPC tumors cluster most similarly to two CNS-PNET molecular subtypes: ependemyoma with *RELA* overexpression (EPN-*REL*) and a subset of CNS-PNETs with no classifiable molecular signature (CNS-PNET-not other specified (-NOS)) (Figure 6A), indicating similar methylation profiles to a subset of primary CNS-PNET patient samples. PFSK-1 cells clustered most similarly to CNS-PNET samples with the high grade glioma-*IDH* mutant (HGG-*IDH*) molecular subtype (Figure 6A). We confirm that the PFSK-1 cells do not have the *IDH* mutation, however a closer look at the methylation profiles reveals that both cohorts exhibit an overall hyper-methylation phenotype compared to other tumor subtypes types, indicating that they may not be systematically or functionally similar at the DNA methylation level. Interestingly, the hIPS and NPC *in vitro* samples clustered separately from the CNS-NPC and PFSK-1 cells *in vivo*, exhibiting more similar methylation profiles to normal human brain tissue (Figure 6A, Supplemental figure 6C). In addition, a progressive methylation phenotype is observed with respect to the NPC-derived tumors, with NPCs *in vitro* exhibiting the most similar methylation patters and IPS cells showing the least similar methylation patterns to CNS-NPCs *in vivo* (Supplemental figure 6B) It is important to note that our clustering method recapitulated similar trends in CNS-PNET patient sample clustering to those that were previously reported, with previously defined molecular sub-types clustering together (Figure 6A)[23, 208]. In addition, t-distributed Stochastic Neighbor Embedding (t-SNE) revealed consistent clustering patterns between the CNS-NPC samples and the EPN-*RELA* subtypes, however the PFSK-1 samples clustered separately from any one subtype, confirming that the hierarchical clustering association between HGG-*IDH* and PFSK-1 tumors may indeed only be driven by a hyper-methylation phenotype, and are not necessarily pathologically similar (Figure 6B).

3.5.2 Molecular Analysis of CNS-PNETs

Previous reports of CNS-PNET primary samples have indicated that specific expression signatures may indicate cell of origin and have predictive potential for tumor behavior and disease prognosis[136].

Specifically, tumors could be classified into one of three potential subtypes that show distinct molecular and behavioral characteristics: primitive neural-like, oligoneural-like, or mesenchymal-like tumors.

Encouraged by the similarities between the CNS-NPC tumors and primary CNS-PNETs at the methylation level, we wondered if the CNS-NPCs may also show the same degree of similarity to primary patient samples at the expression level with regard to their cell of origin. To this end, we performed RNA sequencing analysis on CNS-NPC and PFSK-1 tumors. To compare the expression microarray data of CNS-PNET patient samples from (Sturm et al., 2016) (n=52)[208] with the RNA sequencing data of CNS-NPCs and PFSK-1, we applied the afore-mentioned expression signatures derived in Picard et al., 2012[136] to the primary CNS-PNET patient sample cohort as well as our two CNS-PNET models: CNS-NPC and PFSK-1.

First, hierarchical clustering of primary CNS-PNET patient samples reveals that, on average, patient samples tend to up-regulate the expression of one of the three signatures (Figure 6C). However, with the exception of the CNS-PNET-ETMR molecular subtype, no other molecular subtype showed significant correlation to one particular expression signature, indicating that the cell of origin may not be driving DNA methylation patterning in these tumors.

Hierarchical clustering of CNS-NPC and PFSK-1 tumors reveals markedly different expression patterning between the two models. CNS-NPC tumors exhibit a significant up-regulation of the oligo-neural expression signature (signature 2), most notably characterized by high expression of the oligodendrocyte transcription factors, *OLIG1/2* (Figure 6D). Conversely, PFSK-1 tumors exhibit expression patterns most closely related to the mesenchymal signature, characterized by high expression of the mesenchymal markers, like the twist basic helix-loop-helix transcription factor 2 (*TWST1*) and snail family transcriptional repressor 1 (*SNAI2*) (signature 3) (Figure 6D). We validated these expression patterns at the mRNA level by RT-PCR. Indeed, CNS-NPC tumors exhibited increased expression of *OLIG1/2* relative to human IPS cells (figure 6E and 6F), and more variable expression relative to an mRNA of healthy adult brain tissue (Figure 6H and 6I). By histological analysis of OLIG2, indeed there is some heterogeneity observed in the protein expression of the CNS-NPC tumors, both within and between samples, with approximately 40-60% of cells staining positive (Supplemental figure 6D, 6E and 6F). In addition, CNS-NPC tumors were negative for the mesenchymal marker, *TWST1* (Figure 6G and 6J), and were negative for LIN28 protein expression (Supplemental figure 6D, 6E, and 6F). Conversely, PFSK-1 tumors exhibit low mRNA expression of *OLIG1/2* (Figure 6H and 6I), and high mRNA expression of the mesenchymal marker, *TWST1* (Figure 6G and 6J). In addition, PFSK-1 tumors had no OLIG2 expression and low LIN28 expression at the protein level (Supplemental Figure 6D, 6E, and 6F). In conclusion, through mRNA and protein expression analysis, we identified clear distinctions between our two models, indicating PFSK-1 and CNS-NPCs recapitulate two distinct subtypes of CNS-PNET tumors.

In conducting a thorough assessment and characterization of the methylation, expression patterns of both CNS-NPC and PFSK-1 tumors, we conclude that these models represent two distinct subtypes of CNS-PNET tumors, which recapitulate key features of clinical pathology and molecular biology.

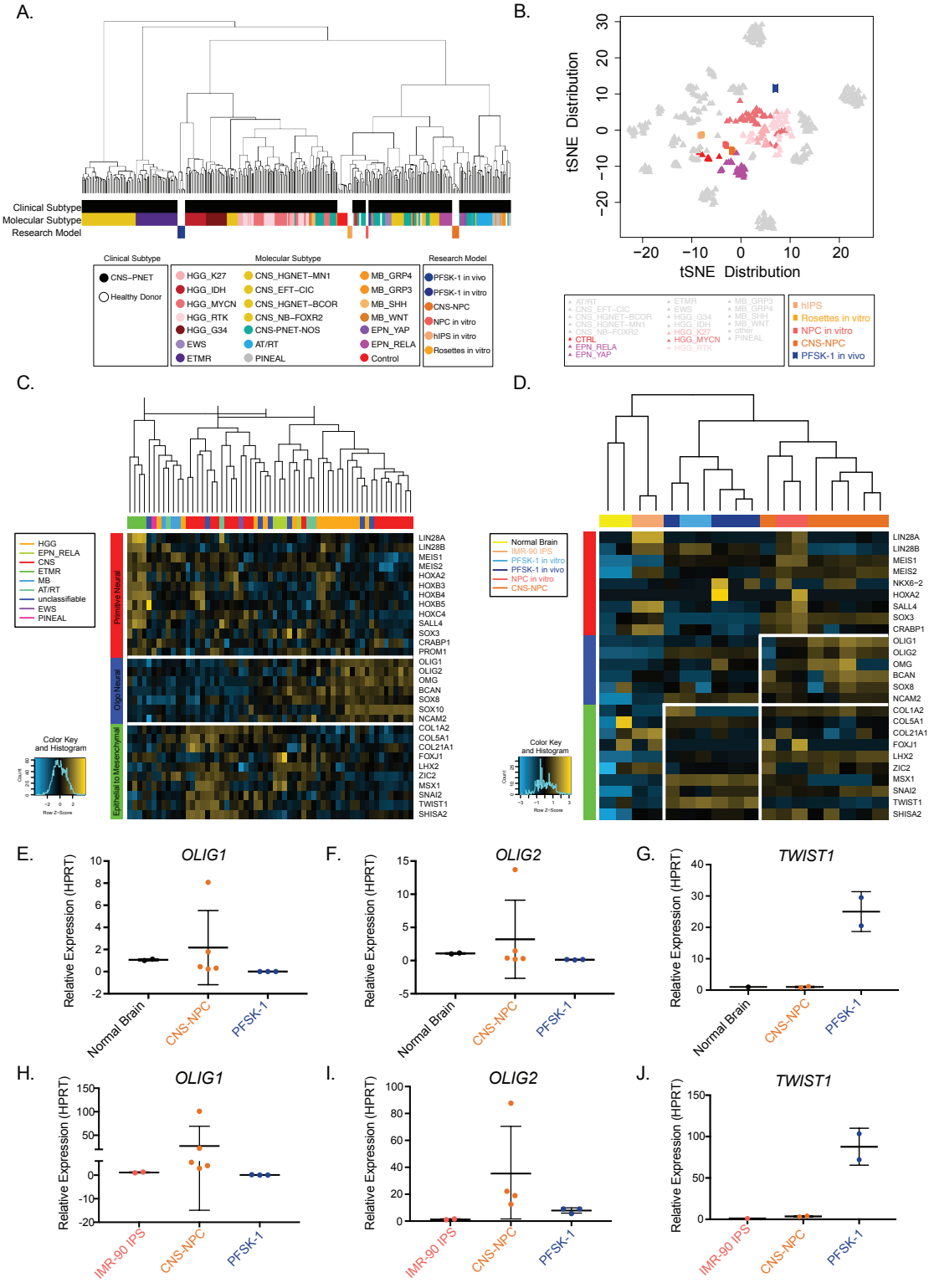
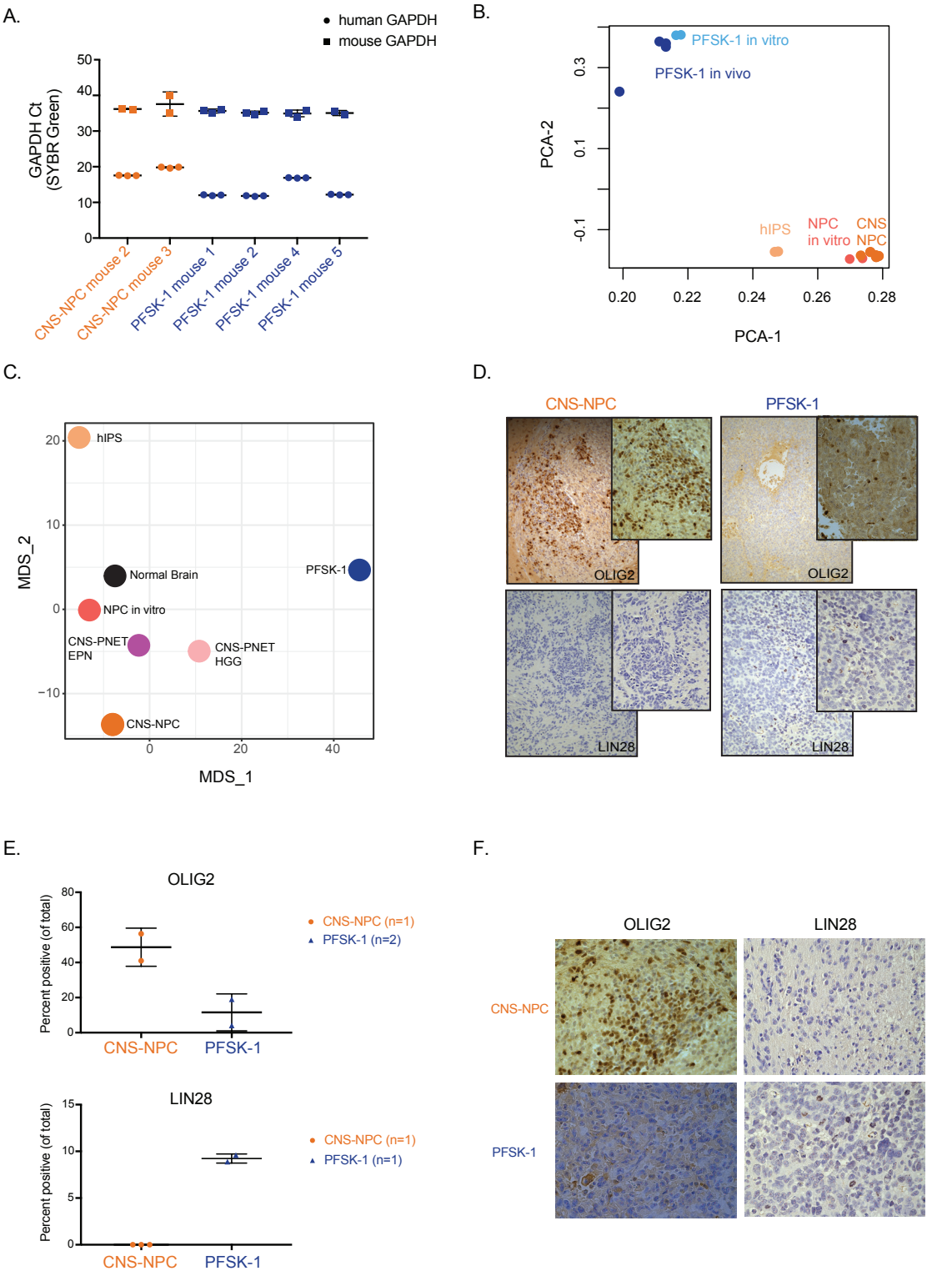


Figure 6. Analysis of DNA Methylation and RNA Expression in CNS-PNETs

(A) Hierarchical clustering of methylation patterns observed in primary CNS-PNET patient samples (n=323), normal brain (n=8), and PFSK-1 (n=5) and CNS-NPC tumors in vivo (n=6) and hIPS and NPCs in vitro (n=2) of the 15,000 most variable probes. (B) T-distributed Stochastic Neighbor Embedding (t-SNE) of methylation patterns observed in primary CNS-PNET patient samples, normal brain, and PFSK-1 and CNS-NPC tumors in vivo and in vitro considering the 15,000 most variable probes. (C) Expression array of three neural cell of origin expression signatures applied to primary CNS-PNET patient samples (n=52) (D) RNA sequencing expression analysis of three neural cell of origin expression signatures applied to PFSK-1 (n=5) and CNS-NPC tumors in vivo (n=6), and PFSK-1, hIPS and NPCs in vitro (n=2), and normal brain tissue (n=2). (E-G) *OLIG1, OLIG2,* and *TWIST1* mRNA expression observed in CNS-NPC and PFSK-1 tumors relative to mRNA expression of normal adult brain, and normalized to *HPRT* mRNA expression. (H-J) *OLIG1, OLIG2,* and *TWIST1* mRNA expression observed in CNS-NPC and PFSK-1 tumors relative to mRNA expression of IMR-90 human IPS cells, and normalized to *HPRT* mRNA expression.



Supplemental Figure 6. Analysis of DNA Methylation and RNA Expression in CNS-PNETs

(A) Cycle quantification (C_q) of *GAPDH* mRNA expression observed in CNS-NPC and PFSK-1 tumors following tissue dissociation and human cell isolation. (B) Pricipal component analysis (PCA) of DNA methylation patterns observed in PFSK-1, IMR-90-hIPS and IMR-90-NPCs in vitro, as well as PFSK-1 and IMR-90-derived CNS-NPCs in vivo. (C) Multidimensional scaling (MDS) of methylation patterns observed in hIPS, NPC in vitro, and PFSK-1 and CNS-NPC in vivo compared with primary CNS-PNET tumor subtypes: CNS-PNET-EPN, CNS-PNET-HGG, and Normal Brain tissue. (D) Representative histology images (10X and 40X magnification) of CNS-NPC and PFSK-1 tumors stained with antibodies targeting OLIG2 and LIN28a/b. (E) Histology quantification of percent (%) of cells positive for OLIG2 and LIN28 expression observed in CNS-NPC and PFSK-1 tumors. (F) Representative image quantification counts for CNS-NPC and PFSK-1 tumors stained for OLIG2 and LIN28.

3.6 Results: Metabolic Characterization of CNS-NPC and PFSK-1 CNS-PNET Tumors

3.6.1 Targeted Mass Spectrometry

Tumor cells exhibit metabolic transformation to support demands of increased proliferation, survival, and growth[65]. Previous studies have shown that tumors with molecular and histological heterogeneity can have different metabolic profiles associated to cell of origin, cell differentiation, and tumor aggressiveness. Given the limitations in diagnosing and monitoring response to treatment using MRI, we wondered if would be possible to exploit differential metabolic signaling to aid in CNS-PNET diagnosis using the CNS-NPC and PFSK-1 tumor models.

Glucose is the key source of energy in the human body and is the primary energy source in the brain. During glycolysis, glucose is oxidized and broken down to produce energy in the form of adenosine triphosphate (ATP). Tumors exhibit hyperactive glucose metabolism to fuel the increased energy demands required to maintain increased proliferation, manage rapid cell turnover, and promote cell survival. To provide a global view of metabolic signaling dependencies in the context of increased energy production needs in CNS-PNETs, we performed an unbiased analysis of glucose metabolism in CNS-NPC and PFSK-1 tumors *ex vivo*. Briefly, brain tissue was isolated from the tumor-bearing hemisphere and the matched contralateral tumor-free hemisphere in NOD-SCID mice intracranially injected with either PFSK-1 of neural progenitor cells (CNS-NPCs) (n=4 for each group). Protein was subsequently purified and extracted for downstream analysis of glucose metabolism using hydrophilic interaction liquid chromatography coupled to tandem mass spectrometry (HILIC - MS/MS) in both positive and negative ionization modes as previously described[186, 198] (Figure 7A). Raw LC-MS data was processed for signal intensity drift correction and noise filtering, where peaks detected at >30% were retained for further analysis.

In an unbiased analysis of glucose metabolism, each tumor exhibited a number of altered metabolites compared to the matched normal contralateral brain tissue (Figure 7B). In a consensus list of 96 metabolites measured by LC-MS, PFSK-1 tumors exhibited significant changes in 49 metabolites ($p<0.05$) (49/96; 51%), and CNS-NPC tumors exhibited significant changes in 43 metabolites (43/96; 45%) (Table 4.4, and Table 4.5). Surprisingly, from the list of significantly altered metabolites, 34 metabolites were shared between the two groups. We conducted a metabolite pathway enrichment analysis of the most significantly altered metabolites ($p<0.05$) in PFSK-1 and CNS-NPC tumors compared with the matched contralateral healthy tissue (Figure 7C and 7D). We identified a number of alterations in cancer-associated pathways, including nucleotide biosynthesis, glutamate metabolism, and CNS-related metabolic processes.

3.6.1.1 Nucleotide Biosynthesis

Both PFSK-1 and CNS-NPC tumors showed the highest enrichment in purine metabolism (CNS-NPC p= 1.22e-5, FDR= 0.00119; PFSK-1 p= 0.00147 FDR= 0.17), as well as other pathways involved in nucleotide biosynthesis, including pyrimidine metabolism and metabolism of nucleotide sugars. 39.5% (17/43) and 26.5% (13/49) of the significantly altered metabolites measured in CNS-NPCs and PFSK-

1 tumors, respectively, are directly involved in nucleotide biosynthetic processes. Indeed, many of the critical metabolic intermediates of the purine and pyrimidine pathways were deregulated in both tumor models, including adenine, guanine, uridine, and cytidine, among others (Figure 7E; Table 4.4 and Table 4.5).

Additional alterations were detected in metabolites that are linked to the increased energetic requirements of tumor cells to support nucleic acid, amino acid, protein biosynthesis, and subsequent cell growth. Aspartate and glycine can contribute to such anabolic processes. First, aspartate acts as a nitrogen donor in the synthesis of inosine, the precursor of purine bases. CNS-NPC and PFSK-1 tumors both exhibited significant decreases in aspartate ($p=0.04$ and $p=0.012$, respectively) (Figure 7E). Inosine was also significantly decreased in CNS-NPC tumors ($p=0.04$). A trend in decreasing levels of inosine was also observed in PFSK-1 tumors, although not significant ($p=0.15$) (Table 4.4 and Table 4.5). Glycine has been previously linked to nucleotide biosynthesis and DNA methylation and is frequently upregulated in cancer. Glycine is increased in both tumor models ($p=0.06$ in CNS-NPCs, and $p=0.0002$ in PFSK-1) (Figure 7E).

3.6.1.2 Glutatmate Metabolism and CNS-Neuraltransmitters

The second most altered metabolic pathway in both PFSK-1 and CNS-NPC tumors is glutamate metabolism. Glutamate metabolism plays a critical role in the central nervous system, synthesizing the excitatory neurotransmitter, N-acetyl-L-aspartyl-L-glutamate, and the inhibitory neuro-transmitter, γ-amino butyric acid (GABA)[193]. While glutamate was not detected in PFSK-1 tumors, glutamate levels were significantly decreased in CNS-NPC tumors ($p=0.035$) (Table 4.4). Interestingly, among other glutamate related metabolites, GABA was also consistently decreased in both CNS-NPCs and PFSK-1 tumors (Figure 7F), consistent with previous observations in neuroblastoma and glioma tumors[37].

3.6.1.3 Glycolysis versus Oxidative Phosphorylation

One hallmark of cancer cells is the up-regulation of metabolic processes to serve the energy requirements for high cellular proliferation and cell turnover[65]. Many, if not all, cancer cells exhibit altered glucose metabolism, where glucose is broken down to produce glucose-6-phosphate, which can be further metabolized to produce pyruvate or can be processed through the pentose phosphate pathway (PPP). Pyruvate supplies energy to the cell either by lactic acid fermentation (aerobic glycolysis) or by entering the mitochondria to begin the citric acid (TCA) cycle. To this end, we wondered which pathways CNS-NPC and PFSK-1 tumors use as their primary source of energy. For each tumor, the endogenous metabolite concentrations are compared to the matched contralateral healthy brain region within the same sample. All results are compiled into one illustrated schematic (Figure 8).

Both CNS-NPC and PFSK-1 tumors exhibited alterations in the metabolites involved in the initial steps of glycolysis, specifically the breakdown of glucose to produce pyruvate. Globally, both CNS-NPC and PFSK-1 tumors exhibit an increase in glycolysis-associated metabolites, including 2,3phopho-glycerate and phosphoenol-pyruvate. Next, while both tumors exhibited an increase in phosphoenol-pyruvate, the precursor of pyruvate, PFSK-1 tumors exhibited a significant increase in lactate ($p=0.01$), but was unchanged in CNS-NPC tumors. Pyruvate can be converted into lactate via lactate dehydrogenase.

Increased levels of lactate are thought to be a direct result of increased aerobic glycolysis in tumor cells[73]. An increase in lactate indicates that PFSK-1 tumors may use aerobic glycolysis for energy production, while CNS-NPC tumors do not.

Alternatively, pyruvate can be transported to the mitochondria to fuel the TCA cycle, where pyruvate is converted to acetyl-CoA, a direct precursor to the TCA cycle. Glutamate and aspartate are also precursors of the TCA cycle, where glutamate is converted to alpha-keto glutarate, a key intermediate in the TCA cycle. Aspartate serves as an electron transporter in the mitochondria to feed the TCA cycle (known as the Malate-Aspartate shuttle), where malate and its intermediate, oxaloacetate, are also key intermediates in the TCA cycle. Interestingly, acetyl-CoA, glutamate, and aspartate were all significantly decreased in CNS-NPC tumors, however, no significant alterations were observed in the metabolic intermediates of the TCA cycle.

Finally, glucose-6-phosphate can also be converted to ribose-5-phosphate (R5P) to be used in the synthesis of nucleotides and nucleic acids. Interestingly, CNS-NPC tumors exhibited a significant decrease in R5P, while PFSK-1 tumors exhibited a significant increase in R5P. These results indicate that there may be some reliance on PPP for nucleotide biosynthesis in PFSK-1 cells, but not CNS-NPC cells.

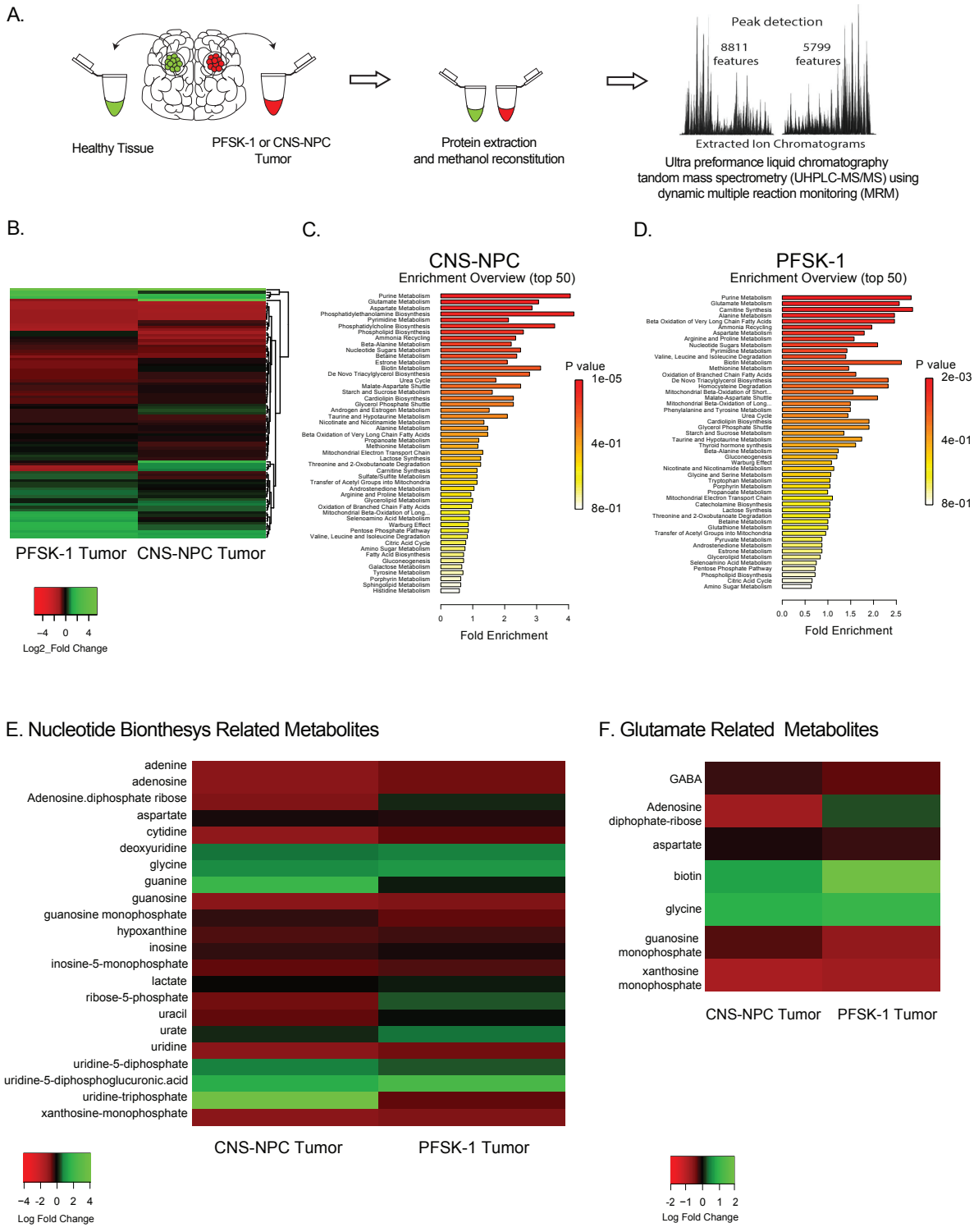


Figure 7. Ex vivo analysis of Glucose Metabolism
(A) Schematic of LC-MS procedure, concentration of metabolites are calculated by the abundance of total ion counts. (B) Hierarchical clustering of the Log2FC of all metabolites measured in both CNS-NPC and PFSK-1 tumors compared with the matched contralateral healthy brain tissue. (C) Metabolic pathway enrichment analysis of significantly altered metabolites in CNS-NPC tumors ($p<0.05$). (D) Metabolic pathway enrichment analysis of significantly altered metabolites in PFSK-1 tumors ($p<0.05$). (E) Log2FC of metabolites involved in nucleotide biosynthesis pathways (F) Log2FC of metabolites involved in glutamate metabolism. Log2FC values were calculated relative to the matched contralateral healthy tissue.

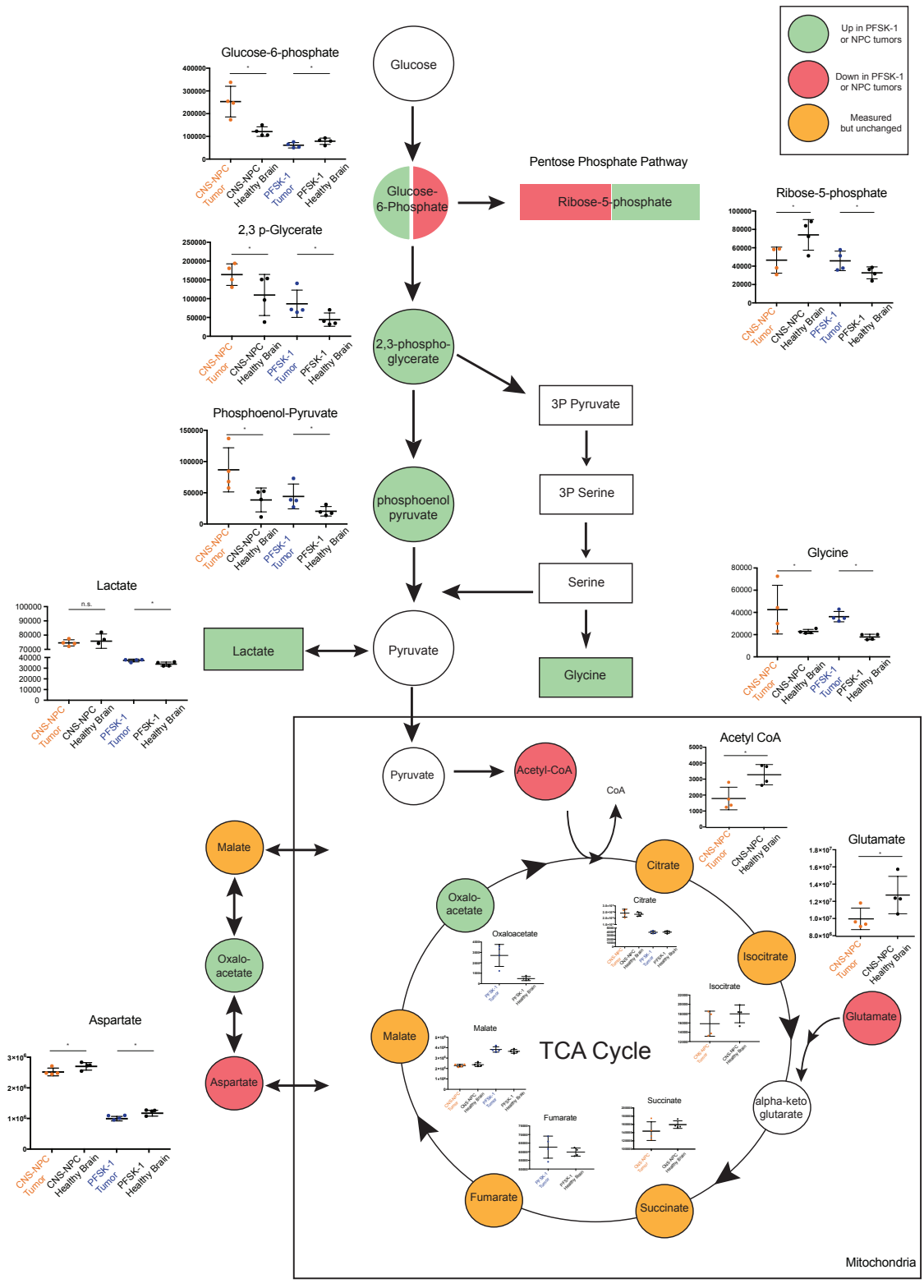


Figure 8. Glycolysis as measure in CNS-NPC and PFSK-1 Tumors
Schematic of glucose metabolism and metabolic alterations observed in CNS-NPC and PFSK-1 tumors compared to the matched contralateral normal brain tissue. Significantly high metabolite concentrations in either PFSK-1 or CNS-NPC tumors are highlighted in green. Significantly low metabolite concentrations in either PFSK-1 or CNS-NPC tumors are highlighted in red. Unchanging metabolite concentrations are highlighted in yellow. Metabolites not measured using LC/MS are not highlighted.

3.7 Results: In vivo metabolic analysis of CNS-NPC and PFSK-1 Tumors

Due to the high number of metabolic alterations observed by LC-MS and given the limitations in diagnosis and monitoring response to treatment implicit to the location of CNS-PNETs, as well as other CNS tumors, we wondered if we could exploit some of the unique metabolic signaling dependencies to identify potential metabolic biomarkers of CNS-PNETS for non-invasive imaging approaches. While some metabolic imaging platforms, like FDG-PET, are not feasible methods for pediatric brain imaging, two techniques have emerged in recent years as reliable methods for tumor detection in the brain: ^{1}H Nuclear Magnetic Resonance and ^{13}C Hyperpolarized Magnetic Resonance.

3.7.1 ^{1}H Nuclear Magnetic Resonance

MRI and ^{1}H neurochemical measurements were acquired in mice bearing CNS-NPC tumors (n=7), PFSK-1 tumors (n=10), as well as mice bearing no tumors, but were stereotaxically injected with buffer (termed: sham control) (n=5) (full spectral profiles of individual measurements are shown in Table 4.6, 4.7, 4.8, and 4.9). Sham control experiments were performed to exclude the possibility of blood brain barrier (BBB) leakage and subsequent aberrant metabolic detection caused by the stereotaxic procedure itself.

First, tumors were detected *in vivo* using magnetic resonance imaging (MRI), which allows for precision mapping of tumor location and size approximation of each mouse evaluated (Supplemental figure 9A and 9F). Following tumor detection, T_2W images guide the placement of the ^{1}H-quadrature surface coil, allowing acquisition of two 2x2x2 mm voxels, one located in tumor area and the other located in the contralateral brain hemisphere. Neurochemical profiles were acquired for each voxel, each representing the endogenous metabolites located in the tumor and the contralateral normal brain tissue (Figure 9A).

Analysis of the neurochemical profiles of each tumor revealed a number of alterations in endogenous metabolite concentrations. Of the 18 different metabolites measured, five metabolites were significantly altered in CNS-NPC tumors compared to the matched contralateral tissue, and four metabolites were significantly altered compared to the sham control tumors (Table 4.6 and 4.9). For PFSK-1 tumors, 10 metabolites were significantly altered compared to the matched contralateral tissue, and eight metabolites were significantly altered compared to the sham control tumors ($p<0.05$) (Table 4.7, 4.8 and 4.9). Measurements of high molecular weight macromolecules (MM) were applied to each image acquisition for baseline peak detection correction (MM-BS)[72]. Importantly, in some brain tumors, MM concentrations can be elevated due to high lipid droplet content[128], however MM concentrations in our datasets were relatively consistent across all acquisition fields, with an average concentration of 1.943 (+/-stdev=0.0409) (Supplemental figure 9C). For all individual metabolites analyzed, the standard error of the peak measured was taken into account. Any outlier values with a standard error of 30% or higher indicates unreliable peak detection, and was removed from further analysis.

Of the significantly altered metabolites, γ-amino butyric acid (GABA) and N-acetyl aspartate (NAA), two important neural transmitters in the central nervous system, were significantly decreased in CNS-NPC and PFSK-1 tumors (Figure 9B and 9C). Measurements of both NAA and N-acetyl-aspartyl-glutamate

(NAAG) were acquired, where NAA was significantly decreased and NAAG was significantly increased in both tumor models (Figure 9C). NAAG is synthesized from glutamate and NAA, indicating that the loss of NAA may, at least in some part, be attributed to an increase in NAAG biosynthesis. Nonetheless, the total (t)NAA levels (NAA+NAAG) were significantly decreased in both models, indicating an overall loss in NAA-dependent glial-specific neurotransmission. In four out of the 10 PFSK-1 tumors measured, the GABA peak was close to the detection limit of the technology. Nonetheless, a closer look at each individual tumor sample showed a high degree of consistency across all PFSK-1 tumors. In addition, GABA peaks were reliably detected in all other conditions measured, indicating this is representative of the true biology of the tumor and not due to technical failures of the acquisition itself. Furthermore, the decrease in GABA observed in ^{1}H MRS is consistent with observations in the LC-MS analysis (Figure 7E). Loss of NAA and GABA metabolites indicate loss of normal neuronal density. Decreases in both GABA and NAA are consistent with previous reports in primary pediatric and adult brain tumors, and have been proposed as biomarkers for detecting brain tumors *in vivo*[7, 37].

Glycine is important for the biosynthesis of amino acids in the cell. Increasing levels of glycine have been found to be correlated poor prognosis in pediatric brain tumors[6], and has been linked to aberrant DNA methylation patterning in cancer[76]. Indeed, both CNS-NPC and PFSK-1 tumors exhibited increases in glycine concentrations (Figure 9D), showing consistencies with the metabolite changes observed in the LC-MS analysis (Figure 7I). In both models, the concentration of endogenous glycine varied considerably between each mouse measurement, but cannot solely be accounted for by the volume of the tumor (Supplemental figure 9F).

Both tumors also exhibited consistent and significant increases in choline concentrations, with both phospho- and glycerophospho-choline contributing to the total choline (tCho) abundance (Figure 9E). This observation is also consistent with previous observations in primary patients, citing its important contribution to the synthesis of new phospholipids, indicating high cellular density and high cellular turnover[7, 63]. Taurine was also significantly decreased in PFSK-1 tumors ($p=0.0027$), but not CNS-NPC tumors, although there appears to be a trending decrease in this cohort, even if not significant ($p=0.1659$) (Figure 9F).

In addition to consistencies observed between the two tumor models, there were two metabolites that were consistently different. As observed in the LC-MS analysis, PFSK-1 tumors showed a significant increase in lactate ($p=0.0072$), while CNS-NPC tumors showed a significant decrease in lactate ($p=0.0298$) (Figure 9G). Conversely, CNS-NPC tumors showed a notable increase in myo-inositol, although not reaching significance ($p=0.0609$). PFSK-1 tumors showed a significant decrease in myo-inositol levels ($p=0.0221$) (Figure 4H). In a closer look at the variability observed in the myo-inositol levels in individual CNS-NPC tumors, the concentration of myo-inositol is highly correlated to tumor size ($R^2=0.8371$) (Supplemental figure 9G). While both metabolites have been strongly linked to cancer metabolic phenotypes, these differences may indicate different metabolic signaling dependences associated with varying stages of tumor aggressiveness.

Analysis of the BBB in each tumor using a gadolinium-contrast agent showed that PFSK-1 tumors consistently disrupted the integrity of the BBB, while the BBB of mice bearing CNS-NPC tumors remained intact (Supplemental figure 9B). Analysis of the sham control also shows an intact BBB, indicating that the disruption observed in PFSK-1 mice was not due to the injection itself, but is inherent to the biology of the tumor. Blood brain barrier disruption can lead to dysregulation of molecules flowing in and out of the brain and may, in part, contribute to some of the metabolic alterations observed in these tumors, particularly myo-inositol and taurine, which both function in osmotic regulation.

3.7.2 ^{13}C Hyperpolarized Magnetic Resonance

Due to the alterations observed in lactate using LC/MS and ^{1}H MRS, we employed ^{13}C hyperpolarized magnetic resonance (HMR) imaging using thermally polarized ^{13}C-pyruvate to measure the enzymatic conversion of pyruvate and lactate *in vivo*. ^{13}C-pyruvate HMR measurements were acquired in CNS-NPC mice (n=4) and PFSK-1 mice (n=3), with separate acquisitions for the tumor bearing hemisphere and the healthy contralateral hemisphere of the brain (Figure 9I). Using ^{13}C-pyruvate HMR, we identified a high shift in the conversion of pyruvate to lactate in PFSK-1 tumors, that were not observed in the matched normal contralateral hemisphere, or in the CNS-NPC animals (Figure 9J). Indeed, CNS-NPC tumors on average exhibited a reduction in the enzymatic conversion from pyruvate to lactate. These results confirm that the increased lactate concentration measured in PFSK-1 tumors using LC/MS and ^{1}H MRS are directly related to the enzymatic activity of these substrates and directly contribute to the final step of aerobic glycolysis. In summary, based on ^{13}C HMR experiments, combined with ^{1}H NMR and mass spectrometry analysis, we conclude the PFSK-1 tumors rely on aerobic glycolysis as a main energy source for fuel in these tumors, while CNS-NPC tumors rely on alternatuve mechanisms (i.e. oxidative phosphorylation) for energy production.

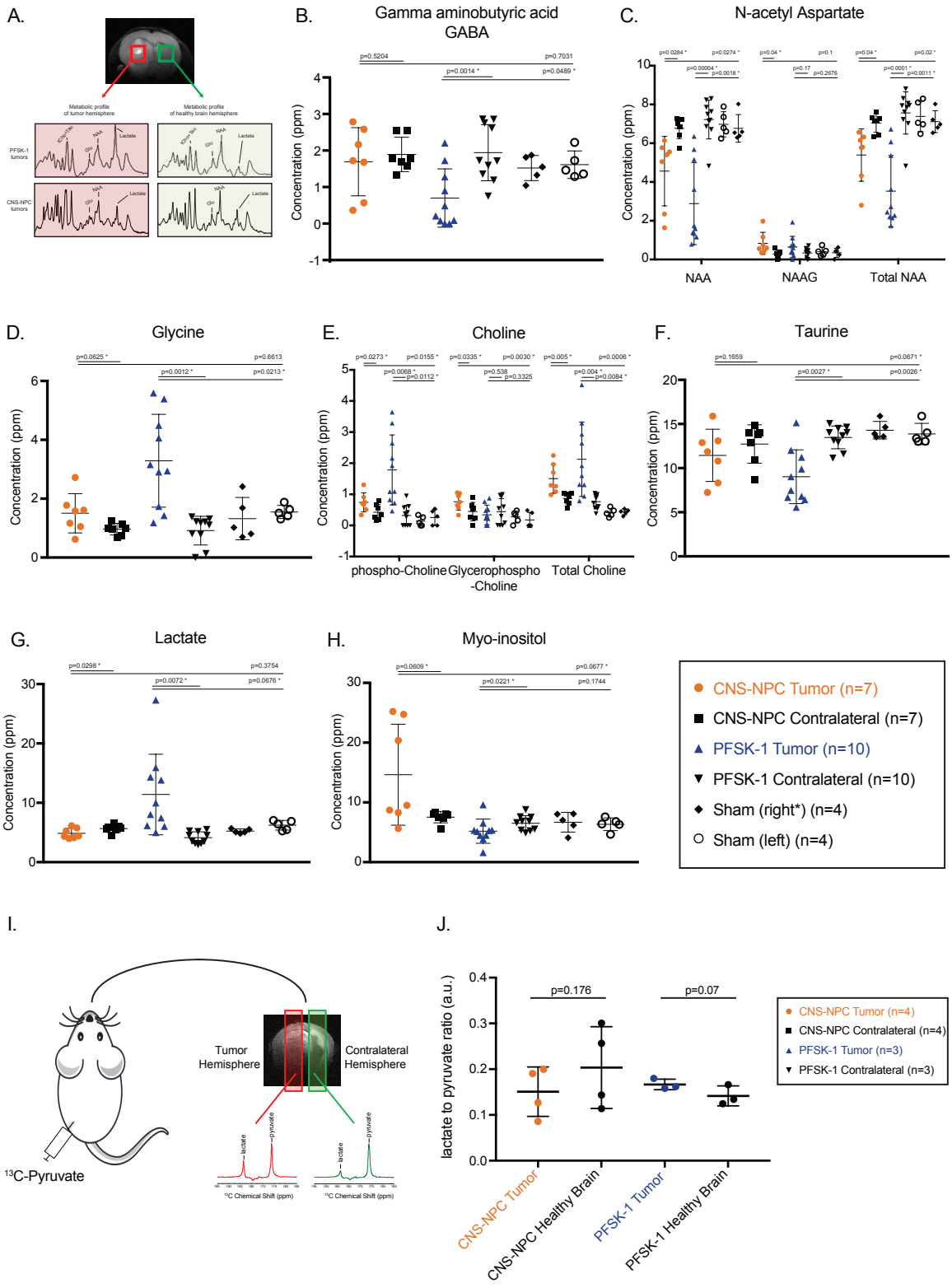
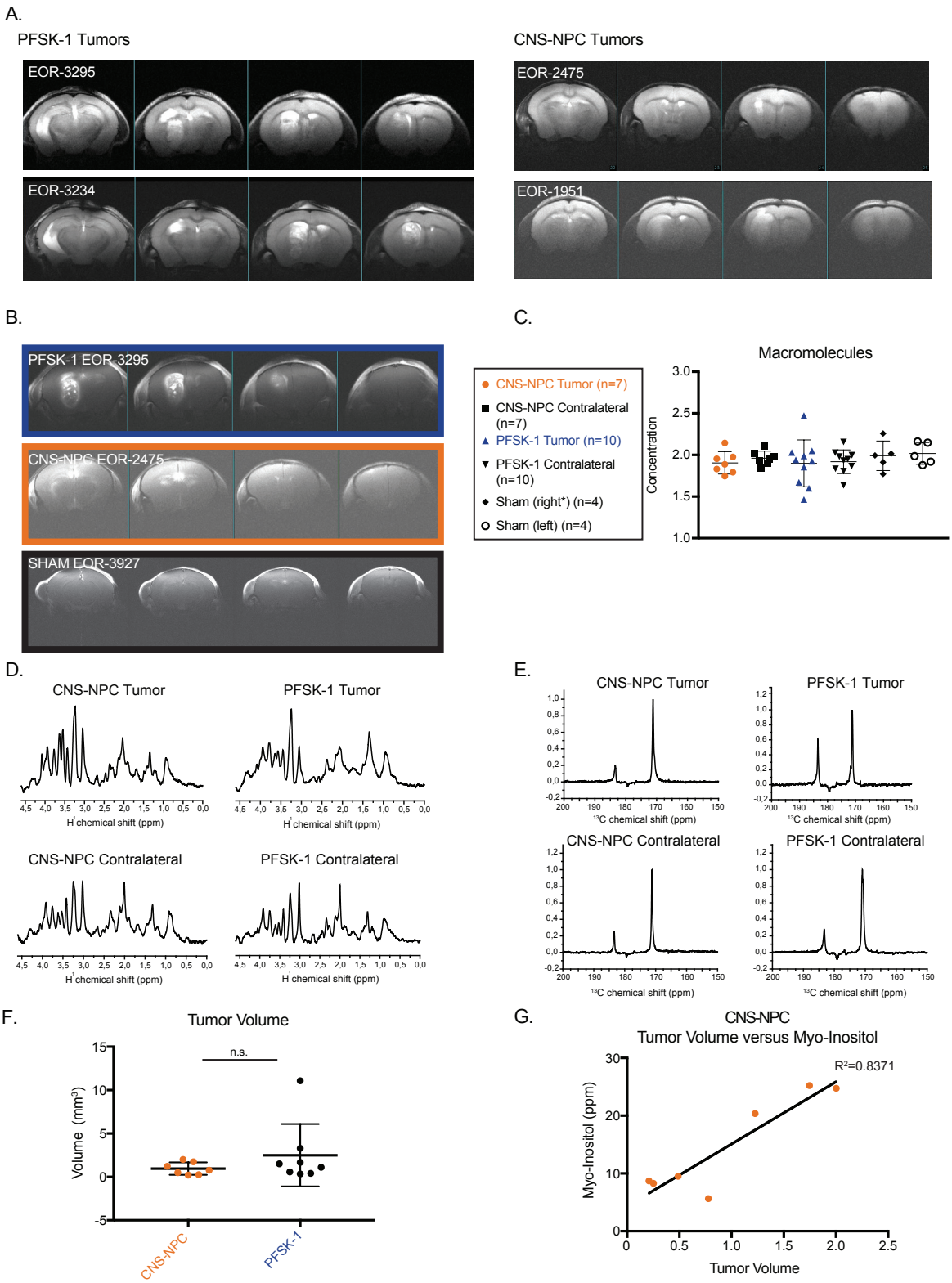


Figure 9. *In vivo* Metabolic Characterization of CNS-NPC and PFSK-1 Tumors
For all individual metabolites: CNS-NPC tumor measurements are in orange, PFSK-1 tumor measurements are in blue, and all control measurements are in black. (A) Schematic of ^{1}H nuclear magnetic resonance imaging (NMR) acquisition. (B-H) Individual metabolite measurements using ^{1}H NMR (n= 7 for CNS-NPC tumors and contralateral hemisphere; n= 10 for PFSK-1 tumors and contralateral hemisphere; n=4 for sham control injection) (I) Schematic of ^{13}C hyperpolarized magnetic resonance (HMR) imaging acquisition. (J) Ratio of pyruvate to lactate peak intensity measured using ^{13}C HMR (n= 4 for CNS-NPC tumors and contralateral hemisphere; n= 3 for PFSK-1 tumors and contralateral hemisphere).



Supplemental Figure 9. ***In vivo*** **Metabolic Characterization of CNS-NPC and PFSK-1 Tumors**
(A) Representative MRI images of the mouse brain bearing PFSK-1 and CNS-NPC tumors. (B) Representative T_2W images + GD-contrast agent for blood brain barrier integrity assessment. (C) Measurement of macromolecules using ^{1}H nuclear magnetic resonance (NMR) (n=7 for CNS-NPC tumors and contralateral hemisphere; n=10 for PFSK-1 tumors and contralateral hemisphere; n=4 for Sham control injection). (D) Representative ^{1}H chemical shifts observed in the mouse brain bearing either CNS-NPC or PFSK-1 tumors. (E) Representative ^{13}C chemical shifts observed in the mouse brain bearing either CNS-NPC or PFSK-1 tumors. (F) Tumor volume of individual CNS-NPC and PFSK-1 mice. (G) Tumor volume versus myo-inositol conc. observed in CNS-NPC tumors.

3.8 Discussion

In this chapter, we explored the epigenetic, transcriptional, and metabolic signaling dependencies of CNS-PNETs using two models of CNS-PNET pathogenesis. In particular, we developed one model, the CNS-NPC model, based on neural progenitor cells derived from human induced pluripotent stem cells isolated and engrafted into immune-comprimised NOD-SCID mice. Using the CNS-NPC model, we demonstrate the ability to generate CNS-PNET tumors derived from a neural-progenitor oligodendrocyte cell of origin. The CNS-NPC model recapitulates many of the pathological, molecular, and metabolic features of a subset of CNS-PNET tumors. In parallel, we use a patient-derived cell line, PFSK-1, as a second model of CNS-PNETs, that also recapitulates features of CNS-PNET tumors, although distinct from the CNS-NPC model. PFSK-1 tumors are likely derived from a more differentiated neuronal cell of origin, and thus exhibit the pathological, molecular, and metabolic features of a more differentiated and aggressive subtype of CNS-PNET tumors.

3.8.1 Development and characterization of two CNS-PNET models

Pediatric primitive neural ectodermal tumors are a rare but aggressive subtype of central nervous system tumor that originate from poorly differentiated or undifferentiated neuro-epithelial cells and often form in the cerebral cortex of the brain. CNS-PNETs are difficult to diagnose and treat due to high molecular heterogeneity and a paucity of defining pathological and genetic features that distinguish this disease. Furthermore, with the exception of one patient derived cell line (PFSK-1), few *in vivo* models exist used to study the unique biology and molecular drivers of these tumors.

In our study, we present a novel *in vivo* model of CNS-PNETs, whereby tumors are generated from neural progenitor cells engrafted into immune compromised (NOD-SCID) mice. Alternative to other stem cell-based models that have been previously presented[118], our model is derived from human induced pluripotent stem cells, which allows for a more accurate mimicry of human biology and disease. We demonstrate that upon engraftment of NPCs, tumor lesions are robustly formed in approximately 2-3 months following stereotaxic injection, termed: CNS-NPC tumors. Histological analyses of CNS-NPC tumors reveal remarkable similarities to primary CNS-PNET tumors, with high expression of neural development and neural stem cell markers, including the neural development type VI intermediate filament protein, NESTIN, and the glial fibrillary acidic protein, GFAP. In addition, CNS-NPC tumor cells exhibit moderate proliferation (40-60%), as measured by Ki-67, and no apoptosis, as measured by the pro-apoptotic marker, CASPASE-III (C3). We demonstrate that the CNS-NPC model is primed by its environment, where subcutaneous engraftment into the flank leads to de-differentiated lesions with low cellular proliferation.

In addition, we analyzed PFSK-1 tumor growth *in vivo*. Contrary to the CNS-NPC tumors, PFSK-1 tumors exhibit highly aggressive behavior, with a three week experimental end-point following engraftment of 10,000 cells. In addition to aggressive growth behavior, PFSK-1 tumors have a high proliferative index (80-100% of cells), and large areas of tumor necrosis. In addition, PFSK-1 tumors had low expression of NESTIN, and do not express GFAP, indicating these tumors may originate from a more differentiated cell state.

Next, CNS-NPCs and PFSK-1 were analyzed at the molecular level through DNA methylation profiling and RNA expression analysis. Comparisons of the DNA methylation profiles and RNA expression profiles with primary CNS-PNET patients revealed that each tumor model exhibits characteristics of distinct subtypes of primary CNS-PNET tumors. At the DNA methylation level, CNS-NPCs share similarities with two molecular subtypes of CNS-PNETS: EPN-*RELA*, and CNS-PNET-NOS. Furthermore, t-distributed Stochastic Neighbor Embedding (t-SNE) further validated the similarities between CNS-NPC tumors and the EPN-*RELA* and CNS-NOS tumor subtypes. Interestingly, the methylation distribution observed by the t-SNE analysis shows a progressive methylation pattern that correlates with different stages of cell differentiation. The most notable changes in DNA methylation patterning are observed upon engraftment, whereby the parental IPS, Rosette, and neural progenitor cells *in vitro* retain DNA methylation features associated to normal brain tissue. Remarkably, however, CNS-NPCs lose methylation patterns associated to normal brain tissue, and gain tumor-associated methylation patterns, pinpointing a unique transformation that occurs in the brain microenvironment. The need for brain-specific priming in the CNS-NPC model is further evidenced by the inability to generate CNS-PNET tumors from NPCs subcutaneously injected into the flank, where proliferation was impaired and neural development markers were no longer retained.

At the DNA methylation level, PFSK-1 tumors exhibit a hyper-methylation phenotype. By unsupervised hierarchical clustering, PFSK-1 tumors showed some similarity to the HGG-*IDH* molecular subtype, but a closer look at the t-distributed Stochastic Neighbor Embedding (t-SNE) of these tumors, PFSK-1 tumors cluster separately from HGG-*IDH*. In addition, PFSK-1 cells do not harbor the *IDH* mutation. Taken together, it is likely that the similarities observed by unsupervised hierarchical clustering are driven by a more generalized hyper-methylation profile, rather than distinct functional methylation events.

RNA expression analysis of early lineage and neural differentiation signatures established by Picard et al., 2012, reveals that the DNA methylation subtypes generally do not correlate to the cell of origin in CNS-PNET patient samples, with the exception of the CNS-PNET-ETMR subtype. CNS-PNET-ETMR tumors clustered exclusively together, exhibiting a high enrichment of genes relating to embryonic stem cells (i.e. the primitive neural signature, #1) compared to other subtypes. ETMR tumors have a hallmark C19MC amplification, which are a cluster of miRNAs that control stemness programs during development and are generally only expressed in undifferentiated or germinal tissues. C19M targets LIN28a, and C19M amplifications directly up-regulate LIN28a to promote tumorigenesis[104, 162]. With regard to the remaining CNS-PNET patient samples, some modest patterning arises with regard to the subtype and the expression of one of the three signatures. Interestingly, a large subset of HGG and unclassifiable CNS-PNET subtypes exhibit an up-regulation of the oligoneural signature (signature 2). While another large portion of unclassifiable CNS-PNETs exhibit an up-regulation of the mesenchymal signature (signature #3).

By applying the same expression signatures to the CNS-NPC and PFSK-1 tumors, we identified distinct expression signatures associated to each of the two models. CNS-NPC tumors exhibit an up-regulation of the oligoneural signature (signature #2), and PFSK-1 tumors exhibit an up-regulation

of the mesenchymal signature (signature 3). Interestingly, at the DNA methylation level, CNS-NPC tumors exhibited similar features to a subset of the HGG and unclassifiable CNS-PNET subtypes. The similarities observed at the DNA methylation level between the CNS-PNET model and patient samples are also observed at the expression level, with both subtypes exhibiting an oligoneural signature, most notably characterized by an increase in the oligoneural development markers: *OLIG1/2*, *BCAN*, and *SOX8*. Further analysis of CNS-NPC tumors confirmed the increased expression of *OLIG1/2* by analysis of mRNA and protein expression, with 40-50% of tumor cells exhibiting elevated OLIG2 expression at the protein level.

Patients with an mesenchymal signature have a high incidence of dissemination and metastasis and exhibit alterations in PTEN and TGF-β signaling[136]. Interestingly, PTEN is a direct target of p53, which is mutated in PFSK-1 cells. Furthermore, it has been shown that mutations in p53 attenuate TGF-β signaling to promote oncogenic programs[35, 81, 83]. Aberrant TGF-β signaling has been shown to regulate tissue interactions[212] and has been linked to tumor cell metastasis[130]. PFSK-1 tumors exhibit rapid proliferation and remarkable tumor mass expansion within the brain. While we did not observe spinal cord metastases in these tumors, the rapid expansion of tumor growth in this particular model (~three weeks) may not lend enough time to observe metastatic spreading and seeding. Further analysis of PFSK-1 tumors confirmed the increased expression of *TWST1* by mRNA expression analysis, and were negative for LIN28 and OLIG2, signature 1 and 2 markers, respectively. While the cell of origin of signature 3 tumors is less clear, given the reduction in expression of genes associated with neural development, it is possible that these tumors arise at a later time during development. This is further evidenced by the age-dependent distribution of the three molecular sub-groups observed in patients, with signature 3 tumors developing more frequently in older children[136]. PFSK-1 tumors were isolated from a patient 10 years of age, which is indeed older than the average age of patients at tumor diagnosis (~5-7 years)[42].

In summary, we developed and characterized a novel *in vivo* model of CNS-PNET pathogenesis based on human neural progenitor cells engrafted into NOD-SCID immune-compromised mice. The CNS-NPC model was developed on the premise that CNS-PNET tumors arise from aberrations that occur during neural development. By isolating and transplanting neural progenitor cells at a precise stage of neuronal differentiation, we are able to mimic these neural development abnormalities in an vivo setting, thereby giving rise to malignant CNS-tumors. Previous studies have shown that a subset of CNS-PNET primary tumors exhibit oligoneural expression signatures, indicating that these tumors may arise from an oligoneural cell of origin. In this study, we demonstrate that oligoneural progenitor cells can indeed give rise to CNS tumors that systematically exhibit features of oligoneural-like CNS-PNETs. The generation of these tumors are not linked to any driving genetic alterations, but are instead driven by the brain micro-environment and epigenetic alterations, specifically DNA methylation.

In addition, by characterizing the CNS-PNET patient-derived cell line, PFSK-1, we performed systematic comparisons of the physiological and molecular features unique to each tumor type. Analysis of both models reveals remarkable heterogeneity between the two tumors at the DNA methylation, RNA expression, and protein level that contribute to the diverse physiological phenotypes observed. CNS-

NPC tumors exhibit a slow onset of tumor development and growth, whereas PFSK-1 tumors are highly aggressive. CNS-NPC tumors exhibit expression patterns linked to early neural development, while PFSK-1 tumors exhibit expression patterns linked to a mesenchymal phenotype and lack expression of neural progenitor and stem cell markers. Finally, CNS-NPC and PFSK-1 tumors exhibit aberrant DNA methylation profiles that correlate to distinct subtypes of CNS tumors. In conclusion, while different from one another, each tumor model captures distinct features, each of which are relevant to CNS-PNET pathogenesis. Furthermore, given the remarkable heterogeneity observed in primary CNS-PNETs, future studies involving both the CNS-NPC and the PFSK-1 models will provide a more complete outlook in understanding the complexities underlying this disease.

3.8.2 LC-MS ex vivo Analysis of Glucose Metabolism

To determine metabolic alterations associated to CNS-PNET pathogenesis, we performed an unbiased analysis of glucose metabolism in CNS-NPC and PFSK-1 tumors. Metabolite concentrations were measured as the total abundance of ion counts present in tissue samples extracted from tumors and the matched contralateral normal healthy brain using liquid chromatography coupled to tandem mass spectrometry. An unbiased analysis of glucose metabolism revealed remarkable similarities in the alterations observed in various metabolic pathways. Specifically, both CNS-NPC and PFSK-1 tumors exhibited alterations in pathways contributing (both directly and indirectly) to nucleotide biosynthesis, CNS neurotransmitters, and DNA methylation. Importantly, alterations in the metabolic processes relating to each of these pathways have been previously linked to cancer and cancer metabolism, particularly in adult and pediatric brain tumors.

PFSK-1 and CNS-NPC Tumors Exhibit Increased Glycolytic Activity

Both CNS-NPC and PFSK-1 tumors exhibited significant increases in the metabolic intermediates involved in the breakdown of glucose, including an increase in 2,3-phosphoglycerate, and phosphoenole pyruvate, the precursor to pyruvate. Increased glucose metabolism to fuel the increased energetic requirements required for rapid cellular proliferation and cell turnover is a hallmark of cancer. The increased concentrations of metabolic intermediates involved in glucose metabolism indicates that both CNS-NPC and PFSK-1 tumors retain this hallmark.

Glycolysis in PFSK-1 tumors

Analysis of all metabolites related to aerobic glycolysis, oxidative phosphorylation, and pentose phosphate pathway (PPP) indicates that PFSK-1 tumors likely rely heavily on aerobic glycolysis for energy production with some activity through PPP. This hypothesis is driven by the fact that PFSK-1 tumors showed significantly high levels in lactate compared to normal tissue. While PFSK-1 tumors showed some alterations in metabolic intermediates of the TCA cycle, many of these metabolites were not detectable using LC/MS. For the few metabolites that were detected, there did not seem to be any consistent alterations to metabolites relating involved in mitochondria-driven oxidative phosphorylation. In addition to aerobic glycolysis, PFSK-1 tumors exhibited an increase in ribose-5-phosphate, a key metabolic intermediate in the PPP pathway, indicating that PFSK-1 tumors may have some reliance on on this pathway for energy production as well.

Glycolysis in CNS-NPC tumors

While there were signifcant alterations observed in early glycolytic intermediates, interestingly CNS-NPC tumors exhibited no consistent alterations in any one of the three pathways analyzed. Specifically, CNS-NPC tumors exhibited no changes in ribose-5-phosphate, indicating limited use of the PPP pathway in this context. In addition, while phosphoenol-pyruvate, the precursor of pyruvate, was significantly increased, Acetyl-CoA and lactate, the two downstream metabolites of pyruvate relating to oxidative phosphorylation and aerobic glycolysis, respectively, were decreased in CNS-NPC tumors. Furthermore, the metabolic precursors of the TCA cycle, aspartate and glutamate, were also significantly decreased in CNS-NPCs, where a decrease in the metabolites feeding into the TCA cycle may indicate high utilization to feed the TCA cycle. However, while no significant alterations were observed TCA cycle intermediates in CNS-NPC tumors compared to the matched contralateral healthy tissue, more TCA-associated metabolites were detectable in CNS-NPC tumors that were not detectable in PFSK-1 tumors.

One critical point is that normal neural tissue was used as a baseline comparison for all metabolites and metabolic pathway analyses. However, it is known that normal neuronal cells sustain a high rate of oxidative phosphorylation to maintain the high energy requirements in the brain, consuming approximately 20% of the circulating glucose in the body[221]. Tumors that rely more heavily on aerobic glycolysis for energy production will have more abundant readout in this particular analysis because neurons don't typically generate lactate for energy production. However, if neurons rely heavily on oxidative phosphorylation for energy production, then the baseline measurement used to compare the reliance on one particular pathway versus another is inherently flawed. Nonetheless, tumor cells that rely on oxidative phosphorylation in the brain still typically exhibit observable changes to metabolites relating to this pathway, so this cannot completely account for the discrepancy.

One additional caveat is presented in the methodology used to extract the tumor tissues. It is likely that the extracted tumor tissues also contain contaminating normal tissue from the surrounding area, leading to an under-sampling of tumor cells in this analysis. As PFSK-1 tumors tend to be, on average, larger than the CNS-NPC tumors, it is likely that there is more contaminating normal tissue in the CNS-NPC tumors. In summary, while it is not possible to conclude reliance on any once single pathway in CNS-NPC tumors from this analysis, there are some indications that CNS-NPC tumors utilize mitochondrial-driven oxidative phosphorylation for energy production.

Taken together, in an unbiased analyses of glucose metabolism, we detected a number of metabolic alterations involved in multiple different pathways and cellular processes in both CNS-NPC and PFSK-1 tumors. Many of these alterations are linked to tumor-associated functions, including nucleotide and amino acid biosynthesis relating to cellular proliferation and cellular turnover, as well as neural-specific programming and DNA methylation.

3.8.3 In vivo Metabolic Analysis of PFSK-1 and CNS-NPC Tumors

Using ^{1}H NMR, we further interrogated these metabolic alterations in CNS-PNETs and confirmed a number of alterations associated to tumor metabolism, including alterations metabolites associated to

neural-specific signaling, including γ-amino butyrate (GABA), N-acetylaspartate (NAA), and taurine. Importantly, low concentrations of these three neural metabolites have previously been observed that neuroblastoma, glioma, and meningioma cells[131], and other tumor subtypes[54, 55]. In addition, alterations were observed in metabolites associated to nucleotide, amino acid and phospholipid biosynthesis, namely glycine and choline. Specifically, up-regulation of glycine has been reported in many tumors, including pediatric brain tumors[6], and has also been linked to aberrant DNA methylation in cancer[76]. Choline is a major source of the methyl groups used in the synthesis of phosphatidylcholine and sphingomyelin, two important phospholipids for the generation of the cell membrane. Indeed, choline has been positively correlated with glioma tumor cell density, and has been proposed as a diagnostic marker for brain tumors as well as a for pharmacodynamic treatment response[10, 63, 187].

In addition to the consistencies observed between the two tumor models, some distinctions in metabolic signaling were also made. Namely, PFSK-1 tumors showed a remarkable increase in the endogenous concentration of lactate, whereas lactate was decreased in CNS-NPC tumors. Indeed, using ^{13}C HMR-DNP imaging, the metabolic conversion of pyruvate to lactate was enhanced in PFSK-1 tumors, but was unchanged in CNS-NPC tumors. These results confirm a reliance on aerobic glycolysis for energy production in PFSK-1 tumors, which is not the case in CNS-NPC tumors.

Finally, CNS-NPC tumors exhibited a strong upregulation of myo-inositol, that was not observed in PFSK-1 tumors. Myo-inositol functions as an osmoregulator in the brain. Increases in myo-inositol have previously been linked to low-grade brain tumors, like low grade gliomas and astrocytomas, but is decreased in high grade gliomas and glioblastoma, which tent to exhibit increases in lactate[164]. The relative flux between myo-inositol pinpoints an attractive marker associated to tumor aggressiveness. CNS-NPC tumors can be categorized as a lower grade tumor compared to PFSK-1 tumors given their overall lower proliferative index (50% versus 90%, respectively), and the overall survival of the animals bearing tumors (2-3 months versus 3 weeks, respectively). Indeed, myo-inositol levels can be linked to damages to the blood brain barrier, where CNS-NPC tumors exhibited an in tact barrier and PFSK-1 tumors exhibited barrier disruption. Nonetheless, damages to the blood brain barrier caused by tumor growth and invasion are also associated to tumor aggressiveness.

In conclusion, we have identified a number of metabolic biomarkers associated to CNS-PNET pathogenesis that could serve as biomarkers for diagnosis and monitoring treatment response in patients. Importantly, these biomarkers have been observed in other brain and solid tumors, indicating these alterations are linked to a global mechanism of dysregulation in cancer. Furthermore, we identified two metabolites, lactate and myo-inositol that can be linked to discreet molecular features of two CNS-PNET tumor subtypes, and may be used as a markers associated to tumor aggressiveness.

3.9 Future Directions and Further Considerations

3.9.1 Generation of CNS-NPC tumors using a second human-IPS cell line

In our CNS-NPC tumor model, we use one human IPS cell line derived from IMR-90 fetal lung fibroblasts (ATCC CCL-186). To exclude the possibility of a cell-line specific phenomenon, we have generated CNS-NPC tumors using a second hIPS cell line (ATCC-DYS0100; ATCC). Upon differentiation into NPCs, efficiency was assessed by RT-PCR and immunofluorescence to confirm efficient neural progenitor differentiation (Supplemental Figure 5B and 5B). We have intracranially transplanted the DYS0100-NPCs into immune-deficient NOD-SCID mice, as previously described, and will sacrifice these animals after 10 weeks for complete histological analysis of molecular and physical characteristics. Specifically, we plan to immuno-stain tissues with neural progenitor markers, Nestin and GFAP. In addition, we will asses proliferation and apoptosis using antibodies targeting Ki-67 and CC3, respectively. Finally, we will also stain tissues with hematoxylin and eosin (H&E) to asses the morphological features of the DYS0100-CNS-NPCs. In conclusion, by generating CNS-NPC lesions from two independent human IPS cell lines, we plan demonstrate that the ability to form a CNS-NPC lesion from neural progenitor cells *in vitro* is not a cell-line specific phenomenon, but is a consequence of a fundamental biological transformation that occurs during aberrant differentiation.

3.9.2 Using the premise of the CNS-NPC model to model additional subtypes of CNS-PNETs

In our study, we present a new model of CNS-PNET pathogenesis based on human cells derived from IPS cells. By isolating and engrafting neural progenitor cells into immune-comprimised NOD-SCID mice, we demonstrate the ability to generate a subset of CNS-PNET tumors that arise from an oligodendrocyte cell of origin that is not dependent on oncogenic driving mutations. In future studies, it will be interesting to demonstrate additional CNS-PNET tumor subtypes that arise at other stages of neuronal differentiation or depend on geneitc mutations as oncogenic drivers in other contexts.

For example, neural rosettes represent the one of the earliest stages of neuronal cell differentiation, whereby neural stem cells are engulfed in rosette niches to poise cells for further differentiation[74], thus representing a primitive neural-like phenotype. One hallmark of embryonal tumors of multi-layered rosettes (ETMRs) is a primitive-neural signature, marked by high expression of LIN28a/b. Thus, we hypothesized that it could be possible to generate ETMR-like tumors using human neural rosettes as the cell of origin. In a preliminary experiment, we isolated rosettes and intracranially transplanted these cells into immune-compromised NOD-SCID mice. After 10 weeks, we isolated and analyzed tissues by histological and DNA methylation analyses. Interestingly, the CNS-Rosette tumors recapitulated some of the morpholoigcal and histological features of ETMRs, most notably characterized by a high abundance of embryonal rosettes, and high expression of LIN28a/b. However, by DNA methylation profiling, these tumors did not exhibit similar mathylation landscape observed in primary ETMR tumors. Thus, from this experiment we conclude that while the physical features of ETMR tumors are present, in this case the cell of origin is not sufficient to generate the unique epigenetic and molecular landscape of ETMR tumors. As mentioned previously, ETMR tumors retain a hallmark feature of the C19MC amplification, which targets a micro-cluster of miRNAs: miR-517c and 520g, and has been previously shown to be an important oncogenic target in ETMR tumors[104]. In future studies, it would be interesting

to constitutively express these two miRNAs in CNS-Rosettes *in vivo*. Results from this experiment would tell us if the overexpression of miR-517c and 520g would affect formation and proliferation of tumors in this context. If tumors form *in vivo*, in a similar fashion as the CNS-NPC tumors, the pathological epigenetic and molecular landscape would be compared with primary ETMR tumors for assesment of similarity between the model and primary samples. Based on previous studies, one might expect an even more elvated expression of primitive neural markers, like LIN28a/b[162]. However, it is unclear how the amplification of miR-517c and 520g would alter the global DNA methylation and transcriptomic landscapes, and future experiments will help bring clarity to these unanswered questions.

In a similar line of thinking, there are subtypes of CNS-PNET tumors harbor genetic driving events. For example, of subset of CNS-PNET tumors harbor a chromosomal copy number gain in 2p24.3, containing the MYC-N oncogene, where the presence of this alteration is associated to poor survival in these patients[58]. In future studies, it would also be possible to genetically modify CNS-NPCs *in vivo* to constitutively express MYC-N, or other oncogenic driving events, and assess the relationship such CNS-NPC-derived tumors and primary patient samples harboring such genetic alterations. While still speculative, if successful, further studies using genetic modifications based on the CNS-NPC model would lend the possibility to raciputulate additional CNS-PNET tumors subtypes and promp further investigative studies into the unique biology of CNS-PNETs.

While such experiments would indeed be very informative, there are number of technical limitations that must be considered. First, if hIPS cell lines are genetically modified, the differentiation potentiation for these cells may be impaired due to the genetic manipulation itself. In addition, the selection of genetically modified cells is difficult because many of the chemical selection agents used (puromycin, for example), are toxic to hIPS cells. In addition, the introduction of genetic modifications with a stable fluorophore timed at the correct stage of neuronal differentiation *in vitro*, rosettes or NPCs for example, would also be technically challenging because a high degree of cell death is observed following FACs sorting, limiting the possibility to test multiple contructs in a timely manner. Furthermore, the biological features of rosettes would not longer be retained if cells were mechanically dissociated into the single cell suspension required for FACs sorting analysis. Therefore, the constructs used in genetic manipulation experiments of hIPS cells, and the timing of the genetic event must be carefully considered.

One interesting approach would be to first validate all genetic constructs using the PFSK-1 cell line, which would allow for rapid testing *in vitro*. Next, unmanipulated cells (Rosettes or NPCs) would be injected *in vivo* (as previously described). Once cells start to expand (as monitored by IVIS imaging), a second injection with concentrated viral particles containing the genetic construct of interest into the tumor location would follow. The use of a stable fluorophore expression would allow for the assessment of genetically modified cells that contribute to tumor expansion, as well as the assessment of their relevance in downstream analyses. While this approach would offset the inherent problems associated with genetic manipulation of human IPS cells *in vitro*, there are still a number of technical considerations that would need to be considered. Most critically, the efficient delivery of the virus to the cells of interest, with minimal or no contact and effects on the surrounding mouse tissue. However, while still speculative, if successful, the expansion of the CNS-NPC model to other subtypes of CNS-

PNETs using genetic manipulation would provide a valuable mechanism of studying the unique and heterogenous biology of CNS-PNET tumors.

3.9.3 Targeting Purine Metabolism in CNS-PNETs

The highest abundance of metabolic alterations observed in both tumors models are those directly involved in purine biosynthesis. Interestingly, purine anti-metabolites were one of the earliest chemotherapeutic agents used to treat cancer, many of which are still routinely used in the clinic today[135, 195]. Methotextrate is a purine antimetabolite and is one of the five standard of care chemotherapeutic agents used for the treatment of CNS-PNETs today[176]. In the SKK-HIT'92 clinical trial, methotextrate treatment was introduced, in combination with clophosphamide, vincristine, carboplatin, and etoposide, as an alternative to radiotherapy. Results from this study showed favorable survival and helped prevent radiotherapy-induced cognitive defects[148]. The abundance of alterations observed in purine metabolism in both tumor models provides a link to methotextrate efficacy in the clinic, as well as pinpointing the potential for additional purine antimetabolites for synergistic efficacy in treating CNS-PNET tumors. That being said, altered purine and nucleotide metabolism is a hallmark of many cancers, and is not unique to CNS-PNET alone, and further studies into the efficacy and mechanism of action of purine antimetabolites are needed, where both PFSK-1 and CNS-NPC tumor models would serve as appropriate platforms for preliminary studies in this context.

3.9.4 Further studies of glycolysis in the CNS-NPC model

Finally, in an *ex vivo* metabolic analysis of CNS-NPC tumors, while signficant alterations were observed in the early steps of glycolysis, no signifcant alterations were observed in the two most abundant metabolic pathways contributing to energy production, specifically aerobic glycolysis and oxidative phosphorylation. Indeed, it is possible that technical limitations of the experimental design may be contributing to these confounding results. One alternative way to address this open-ended question experimentally would be to measure glucose metabolism in CNS-NPC tumors using ^{13}C hyperpolarized magnetic resonance (HMR) using glucose as a metabolic tracer. Using ^{13}C-tagged glucose HMR imaging would allow for the complete tracing of glucose breakdown in the CNS-NPC tumors in real time, where the contribution of the metabolic intermediates towards aerobic glycolysis or oxidative phosphorylation could, in theory, be discriminated. Given that the CNS-NPC tumors require the brain niche for transformation, similar metabolic tracing experiments *in vitro* are unlikely to produce results that can be translated to the *in vivo* setting. ^{13}C-HMR analyses offer a unique way of monitoring metabolic processes in real time, in living tissue. While pyruvate tracers in ^{13}C HMR experiments are informative for understanding metabolic dependencies on aerobic glycolysis, glucose as a metabolic tracer offers the opportunity to expand our understanding of the underlying processes responsible for glucose breakdown in different contexts, including both aerobic glycolysis and oxidative phosphorylation.

Chapter 4. Materials and Methods

Dissecting epigenetic, transcriptomic, and metabolic signaling dependencies using a novel model of central nervous system primitive neural ectodermal tumors (CNS-PNETs)

Table 4.1: List of Critical Reagents Used in Chapter 3

Name	Company	Catalogue Number
pHAGE PGK-GFP-IRES-LUC-W	Addgene	46793
EZ DNA Methylation kit	Zymo	D500
RNAeasy Mini Kit	Qiagen	74104
MACS brain tumor dissociation kit	Miltenyi Biotec	130-095-942
MACS Mouse cell depletion kit	Miltenyi Biotec	130-104-694
Matrigel	Corning	356234
mTeSR™ 1 cell culture medium	Stem Cell	85850
Matrigel basement membrane (hESC qualified)	Corning	354227
Stemdiff™ Neural Progenitor Medium	Stem Cell	5833

4.1 Cell culture

PFSK-1 cells were purchased from ATCC (CRL-2060) and authenticated by STR (Microsynth, Switzerland). Cells are propagated in RPMI cell culture medium (Life Technologies) supplemented with 10% fetal bovine serum (Life Technologies) and 1% pen-strep.

Human-IPS (hIPS) cells were generated from IMR-90 fetal lung fibroblasts (ATCC CCL-186) by viral transduction of a combination of OCT4, SOX2, MYC and KLF4 genes[129]. Neural progenitors were generated using a modified dual SMAD-inhibition protocol as previously described[168, 93].

ATCC-DYS0100 human-IPS (hIPS) cell line were purchased from ATCC (ACS-1019).
hIPS cell lines were maintained and propagated in mTeSR™ 1 cell culture medium (Stem Cell) and passaged onto matrigel basement membrane (hESC qualified) coated dishes (Corning). Neural progenitor cells were maintained and propagated in Stemdiff™ Neural Progenitor Medium (Stem Cell).

4.2 Lentiviral Infection

Lentiviral delivery of pHAGE PGK-GFP-IRES-LUC-W (Addgene) to generate cell lines stably expressing luciferase and GFP together. Cells containing the vector were selected based on GFP expression.

4.3 Stereotaxic Injection

Cells were intracranially transplanted into male or female immune-deficient NOD-SCID mice 4-8 weeks of age (Jackson Laboratories). Briefly, cells are targeted into the right lateral ventricle of the cerebral cortex and injections are performed as follows: a 25-guage burr hole is made approximately +0.5 ventral, and 1.8mm right or left lateral from the bregma. A 26-gauge needle attached to a 10 µl Hamilton syringe is inserted into the depth of 3.2mm from the skull surface using stereotactic guidance. Two microliters of 1X PBS containing cells (10,000 cells for PFSK-1; 200,000 cells for IPS-derived rosettes or neural progenitor cells) are inoculated into the brain over a period of five minutes. Cells are left to settle for an additional three minutes prior to needle resection. Tumor growth is monitored weekly or bi-weekly using the IVIS system.

4.4 Flank Injection

Cells were mixed in a 1:1 ratio of 1X PBS containing cells (10,000 cells for PFSK-1; 200,000 cells for IPS-derived rosettes or neural progenitor cells) and Matrigel (Corning). 100µl of the 1:1 mixture were injected subcutaneously into the flank of the mouse, under the skin overlying the upper thigh and lower part of the back. Tumor growth was monitored weekly or bi-weekly using the IVIS system.

4.5 Immunohistochemistry

Tissues are fixed in 4% PFA solution for 4 hours, washed in 1X PBS, and submerged in 70% EtOH prior to paraffin embedding. Briefly, the general morphology of paraffin embedded tissues is assessed on slides stained with Haematoxylin-Eosine (H&E). Immunohistochemical detection was performed using the fully automated Ventana Discovery XT (Roche Diagnostics, Rotkreuz, Switzerland). All steps were performed on the machine with Ventana solutions. Briefly, dewaxed and rehydrated paraffin sections were pretreated with heat using standard condition (30 minutes) CC1. The primary antibodies were incubated 1 hour at 37°C. After incubation with rabbit OmniMap (Roche Diagnostics, Rotkreuz, Switzerland), chromogenic drevelation was performed with ChromoMap kit (Roche Diagnostics, Switzerland).

Table 4.2: List of Antibodies Used

Target Name	Clone	Species	Company	Catalogue Number	Dilution Factor
Nestin	10C2	Rabbit	Abcam	ab22035	1000
GFAP	polyclonal	Goat	Sigma	SAB25000462	400
Olig2	polyclonal	Rabbit	Abcam	ab42453	200
CC3	ASP175	Rabbit	Cell Signaling	9664	200
Ki67	SP6	Rabbit	Abcam	ab21700	200

4.6 Human cell isolation

Animals were sacrificed following Swiss animal regulatory guidelines (Service de la consommation

et des affaires vétérinaires, Canton de Vaud, Suisse). Brain tissue was isolated and enzymatically and mechanically dissociated using the MACS brain tumor dissociation kit (Miltenyi Biotec) followed by human cell isolation using the MACS Mouse cell depletion kit (Miltenyi Biotec). RNA and/or DNA of isolated and purified viable human cells were further analyzed as follows:

DNA Extraction: Genomic DNA was extracted using phenyl/chloroform followed by ethanol (EtOH) precipitation

RNA Extraction: total RNA was extracted from cells by phase separation using tizol and chloroform followed by RNAeasy Mini Kit (Qiagen)

4.7 Quantitative PCR (RT-PCR)

Reverse transcription was preformed using Superscript III first strand synthesis kit (ThermoFisher Scientific). Quantitative PCR (qPCR) is preformed using SYBR Green real time PCR master mix (ThermoFisher Scientific) and primers designed to specifically target genes of interest.

Table 4.3: List of Primers Used for RT-PCR

All oligonucleotides were synthesized by Microsynth, Switzerland

Target Name	Forward Primer (5'-3')	Reverse Primer (5'-3')
HPRT	ACC CTT TCC AAA TCC TCA GC	GTT ATG GCG ACC CGC AG
NESTIN	CAG CGT TGG AAC AGA GGT TGG	TGG CAC AGG TGT CTC AAG GGT AG
GFAP	CAT TCC CGT GCA GAC CTT CT	ACG GTC TTC ACC ACG ATG TT
NEFM	AGG CCC TGA CAG CCA TTA C	CTC TTC GGC TTG GTC TGA CTT
PAX6	TGG GCA GGT ATT ACG AGA CTG	ACT CCC GCT TAT ACT GGG CTA
SOX2	TAC AGC ATG TCC TAC TCG CAG	GAG GAA GAG GTA ACC ACA GGG
OLIG1	GAG GTC ATC CTG CCC TAC TC	CGT GGC TAT CTT GGA GAG CTT
OLIG2	TTG AAG TCA TCC TCG TCC AGC	TCGCGGCTGTTGATCTTGAG
LIN28B	CAT CTC CAT GAT AAA CCG AGA GG	GTT ACC CGT ATT GAC TCA AGG C
TWST1	ATT CAG ACC CTC AAG CTG GC	TCC ATC CTC CAG ACC GAG AA

4.8 DNA Methylation Bisulfite Conversion

Genomic DNA was treated with bisulfite for subsequent DNA methylation profiling using the EZ DNA Methylation kit. DNA methylation profiling was performed using the HymanMethylation850 BeadChip methylation microarray (Infinium).

4.9 DNA Methylation Analysis

The global DNA Methylation profiles of our samples were compared to the landscape of DNA

methylation in CNS-PNET tumors[208]. The raw DNA methylation data for our samples (HymanMethylation850 BeadChip) were processed using the R package minfi (Biodirector, 10.18129/B9.bioc.minfi). Methylation data from [Sturm et al., 2016] were downloaded from [GSE73801]. To remove any batch effect between the two datasets, we co-normalized their beta values matrices using ComBat, after retaining the probes in common between the two datasets. Values rescaled outside the range [0,1] were clipped.We compared the methylation status of our samples to the PNET dataset by means of hierarchical agglomerative clustering, PCA and t-SNE analysis. In the hierarchical clustering, in order to speed up computation and remove background noise, we only considered the 15,000 most variable probes. Additionally, we retained only the probes spanning a wide range of beta values, by removing all the probes without at least one beta value smaller than 0.2 and one greater than 0.8. Results were stable as we lowered the lower numbers of probes considered.

The t-SNE analysis was performed with the R package Rtsne. We checked the robustness of the results at different levels of perplexity and found them to be consistent within a wide range (from 10 to 50). Results were also unchanged between considering the entire set of probes or only the variable ones (as done for the hierarchical clustering approach). In the final results we chose a perplexity of 30 for the embedding.

4.10 RNA expression analysis of patient samples

Purified RNA from CNS-NPC and PFSK-1 tumors were sequenced with the Illumina NextSeq500.

Differential Expression analysis

Expression data from [Sturm et al., 2016] were downloaded from [GSE73038]. Differential expression was assessed using the limma package, P values were adjustedusing the Benjamini–Hochbergmethod and significance cut-offset at 0.01.

Gene Ontology (GO) Terms enrichment analysis

GO Terms enrichment analysis was performed using the online webservice based on the Molecular Signatures Database (MSigDB) (http://software.broadinstitute.org/gsea/msigdb/annotate.jsp). Enriched GO terms were defined as GO biological process (BP) and molecular function (MF) terms obtaining a FDR-adjusted P value <0.01, retreiving a maximum of 100 terms.

Gene expression signature enrichment analysis

Single sample Gene Set Enrichment analysis (Barbie et al., 2009) (ssGSEA), implemented in the R package GSVA, was used to calculate an expression score for each gene expression signature and each sample. The default parameters from the GSVA package were used.

4.11 Mass Spectrometry

Metabolite extraction

Tissue samples were pre-extracted and homogenized1,2 by the addition of 150 µL of MeOH:H2O (4:1) per 10 mg of frozen tissue weight, in the Cryolys Precellys 24 sample Homogenizer (2 x 20 seconds

at 10000 rpm, Bertin Technologies, Rockville, MD , US) with ceramic beads. The bead beater was air-cooled down at a flow rate of 110 L/min at 6 bar. Homogenized extracts were centrifuged for 15 minutes at 4000 g at 4°C (Hermle, Gosheim, Germany). The resulting supernatant was collected and evaporated to dryness in a vacuum concentrator (LabConco, Missouri, US).

Protein quantification

The protein pellets were evaporated and lysed in 20 mM Tris-HCl (pH 7.5), 4M guanidine hydrochloride, 150 mM NaCl, 1 mM Na2EDTA, 1 mM EGTA, 1% Triton, 2.5 mM sodium pyrophosphate, 1 mM beta-glycerophosphate, 1 mM Na3VO4, 1 µg/ml leupeptin using brief probe-sonication (5 pulses x 5 sec). BCA Protein Assay Kit (Thermo Scientific, Masschusetts, US) was used to measure (A562nm) total protein concentration (Hidex, Turku, Finland).

Sample amount normalization

The dry extracts were reconstituted using MeOH:H2O (4:1, v/v) and the volume of re-suspension was adjusted based on measured protein quantities in each pellet.

Data Acquisition- LC-MS Analysis

Extracted samples were analyzed by Hydrophilic Interaction Liquid Chromatography coupled to tandem mass spectrometry (HILIC - MS/MS) in both positive and negative ionization modes using a 6495 triple quadrupole system (QqQ) interfaced with 1290 UHPLC system (Agilent Technologies). In positive mode chromatographic separation was carried out using a BEH Amide, 1.7 µm, 100 mm × 2.1 mm I.D. column (Waters, Massachusetts, US). Mobile phase was composed of A = 20 mM ammonium formate and 0.1 % FA in water and B = 0.1 % FA in ACN. The linear gradient elution from 95% B (0-1.5 min) down to 45% B was applied (1.5 min -17 min), and 45% B was held for 2 min. The initial chromatographic conditions were maintained as a post-run during 5 min for column re-equilibration. The flow rate was 400 µL/min, column temperature 25 °C and sample injection volume 2µl. ESI source conditions were set as follows: dry gas temperature 290 °C, nebulizer 35 psi and flow 14 L/min, sheath gas temperature 350 °C and flow 12 L/min, nozzle voltage 0 V, and capillary voltage 2000 V. Dynamic Multiple Reaction Monitoring (DMRM) was used as acquisition mode with a total cycle time of 600 ms. Optimized collision energies for each metabolite were applied.

In negative mode, a ZIC pHILIC (100 mm, 2.1 mm I.D. and 5 µm particle size) column was used. The mobile phase was composed of A = 20 mM ammonium Acetate and 20 mM NH4OH in water at pH 9.3 and B = 100% ACN. The linear gradient elution from 90% B (0-1.5 min) down to 50% B was applied (1.5 min – 8 min), followed by an isocratic step (8 min – 11 min) and linear gradient down to 45% B (11 min – 12 min). These conditions were held 3 min. Finally, the initial chromatographic conditions were established as a post-run during 9 min for column re-equilibration. The flow rate was 300 µL/min, column temperature 30 °C and sample injection volume 2µl. ESI source conditions were set as follows: dry gas temperature 290 °C and flow 14 L/min, sheath gas temperature 350 °C, nebulizer 45 psi, and flow 12 L/min, nozzle voltage 0 V, and capillary voltage -2000 V. Dynamic Multiple Reaction Monitoring (DMRM) was used as acquisition mode with a total cycle time of 600 ms. Optimized collision energies for each metabolite were applied.

Quality Control

Pooled QC samples (representative of the entire sample set) were analyzed periodically (every 4 samples) throughout the overall analytical run in order to assess the quality of the data, correct the signal intensity drift (attenuation in most cases, that is inherent to LC-MS technique and MS detector due to sample interaction with the instrument over time) and remove the peaks with poor reproducibility (CV > 30%).

Data (Pre) Processing

Raw LC-MS/MS data was processed using the Agilent QqQ-Quant software (version B.07.00, Agilent technologies). The relative quantification of metabolite features was based on EIC (Extracted Ion Chromatogram) areas for the MRM transitions monitored. The obtained metabolite table (containing peak areas of all detected metabolites features across all samples) was exported to "R" software http://cran.r-project.org/ and signal intensity drift correction and noise filtering (CV (QC features) > 30%) was done within the "batchCorr" R package6.

Statistical Analyses (Univariate and Multivariate)

Parametric Student t-test* (or non-parametric Mann-Whitney) was applied to extract the significantly altered metabolite features between PFSK-1 / NPC tumor tissue and contralateral tissue with an arbitrary level of significance, p-value = 0.05 (and q-value as an adjusted p-value using an optimized False Discovery Rate approach). Data were log transformed prior to t-test application. Multivariate statistical analysis were achieved with SIMCA software (Version 14.1, Umetrics, Sweden) and used for data quality assessment and exploration.

4.12 In vivo Metabolic Imaging

Magnetic resonance imaging

All measurements were performed on a Varian INOVA spectrometer (Varian, Palo Alto, CA, USA) interfaced to a 31-cm horizontal-bore actively shielded 9.4T magnet (Magnex Scientific, Abingdon, UK).

Animal preparation

All experimental procedures involving animals were conducted according to federal and local ethical guidelines and the protocols were approved by the local ethics committee of the Canton Vaud, Switzerland (Service de la consommation et des affaires vétérinaires, Canton de Vaud, Suisse). Tumor bearing or sham operated male or female NOD-SCID mice were anesthetized using 1.5 ± 05% isoflurane (Attane, Minrad, NY, USA) in 30% oxygen and air. A femoral vein was catheterized to deliver hyperpolarized pyruvate solution. After surgery, the animal was placed on the MRI bed, the head was fixed using stereotaxic system (RAPID Biomedical Inc., OH, USA). Body temperature was monitored by a nonmagnetic rectal probe and maintained at 37.0 ± 0.5 °C by warming the animal with temperature-controlled water circulation (SA instruments Inc. NY, USA), and their respiration rate was monitored using respiration from a small pneumatic pillow sensor (SA Instruments Ins. Stony Brook, NY, USA). Animal physiology was monitored during the entire duration of the experiment.

Proton MRS

Proton images and spectra were acquired using a home-built ^{1}H-quadrature surface coil that was placed on top of the mouse head. T_2W images were acquired using fast spin echo multi slice (FSEMS) protocol (FOV 18x18 mm2, TR = 4000 ms, effective TE = 52, 6 scans). B0 inhomogeneity was corrected using the FASTESTMAT algorithm[62] in two a 2x2x2 mm voxels, one located at the tumor area and the other located at the contralateral hemisphere. To follow variations in the neurochemical profile single voxel ^{1}H MRS measurements were acquired in those voxels using the SPECIAL pulse sequences[117] (TR = 4000 ms, TE = 2.8 ms, 200 ms acquisition time in 15 x 16 scans). Metabolite concentrations were calculated using LC Model-based fitting routine[139]. The integrity of the blood brain barrier (BBB) was assessed by T_1 weighted (T_1W) images acquired after injection of gadolinium-contrast agent (Gadovist® 5uL/gr body weight) images that were acquired after the completing of the HP ^{13}C MRS measurements, using FSEMS protocol (FOV 18x18 mm2, TR = 350 ms, effective TE = 11 ms, 6 averages)

Hyperpolarized ^{13}C MRS

HP [$^{1\text{-}13}$C]pyruvate solutions were prepared as previously described[196]. Briefly, neat [$^{1\text{-}13}$C] pyruvic acid (Sigma Aldrich, Basel Switzerland) was mixed with trytil radical OX63 (21 mM, Albeda, Denmark). 10 uL of the solution was dynamically polarized for 2 hr using a custom-designed 7Tesla DNP polarizer operating at 197 GHz / 1.00±0.05K5, in order to balance the pH upon dissolution, an additional 13 uL frozen bead of 10 M NaOH solution was added to the sample cup. Once reaching maximal polarization the solution was rapidly dissolved with 5 mL of superheated deuterated phosphate buffer and transferred within 2s into the separator/infusion pump[32], that was prepositioned inside the magnet bore at the vicinity of the animal. A bolus of the solution was automatically infused as previously described[29].

^{13}C MR measurements were performed using a home-built ^{1}H-quadrature/^{13}C-single loop surface coil that was placed on top of the mouse head. B0 inhomogeneity was corrected using FASTESTMAP algorithm[62] in a 3.6x6.9x4.5 mm voxel_ENREF_8. A 325 μL of 25 mM HP [$^{1\text{-}13}$C] pyruvate solution was injected by the automated protocol[29]. A series of slab excitation pulse sequence was then triggered using a slice selective sync pulse with an average flip angle of 30° every 1.5 s and alternating between two 3 mm slabs. The first was positioned at the tumorous hemisphere and the second at the contralateral one. The area under the curve (AUC) of the metabolite was calculated using VNMRJ software by integrating the summed ^{13}C MRS the spectra after phase and baseline correction. The peak areas of [$^{1\text{-}13}$C]pyruvate (172 ppm), and [1-13C]lactate (183.5 ppm) were quantified, and used to compute the lactate-to-pyruvate (LPR) ratio. The HP ^{13}C contrast between the two hemispheres was calculated from the LPR in the tumor slab to the LPR ratio at the contralateral slab.

Table 4.4. List of metabolites measured in CNS-NPC tumors and contralateral tissue using LC/MS

Metabolite	NPC Control Ave	NPC Tumor Ave	Fold Change	Log2 FC	Ttest	NPC Tumor 1	NPC Tumor 2	NPC Tumor 3	NPC Tumor 4	NPC Contra 1	NPC Contra 2	NPC Contra 3	NPC Contra 4
ALANINE	1373664.77	1404360.19	1.02	0.03	0.36	1598457	1450892	1328892	1239200	1417451	1366985	1282882	1427341
ASPARAGINE	17702.66	35129.48	1.98	0.99	0.04	54323	41422	27737	17035	18378	17766	14636	20030
ASPARTATE	2712279.09	2532485.94	0.93	-0.10	0.04	2484361	2449520	2714818	2481245	2836221	2725710	2739265	2547920
GLUTAMATE	12729603.98	9959847.26	0.78	-0.35	0.03	9609747	9076225	9345957	11807460	12288443	10528706	12371879	15729387
GLUTAMINE	34381506.40	36004679.53	1.05	0.07	0.36	36138103	45085329	32278695	30516592	29667413	41266277	28762275	37830061
N-ALPHA-ACETYLLYSINE	64746.84	75376.97	1.16	0.22	0.09	83596	82504	77612	57796	58831	68184	57887	74085
N,N-DIMETHYL ARGININE	915215.99	1470688.22	1.61	0.68	0.03	1765589	1835923	1470022	811219	729632	1153040	818913	959279
N-ACETYL NEURAMINATE	547994.64	418590.89	0.76	-0.39	0.03	421489	392402	318043	542430	538564	529325	494306	629784
URIDINE DIPHOSPHATE GLUCURONIC ACID	12724.62	35369.96	2.78	1.47	0.04	61633	42158	26721	10967	9528	10895	13514	16962
4-ACETAMIDO BUTANOATE	30725.00	119263.51	3.88	1.96	0.00	139659	152201	134380	50814	33764	33976	28671	26488
4-GUANIDINO BUTANOATE	7121448.48	9310834.77	1.31	0.39	0.03	9025068	11837899	9355800	7024571	7089974	7105486	7056220	7234114
4-HYDROXY-PROLINE	1021572.00	1004445.78	0.98	-0.02	0.24	1039582	990142	1010678	977382	1028688	981586	1008912	1067101
CREATININE	8224004.27	9025166.76	1.10	0.13	0.18	8829732	10175478	8657542	8437914	6407424	7773658	9060590	9654345
GAMMA-AMINOBUTYRATE	3041633.00	2628200.51	0.86	-0.21	0.04	2560209	2825419	2568269	2558904	3497051	3038329	3025975	2605178
N-ACETYL PUTRESCINE	7521.00	217172.40	28.88	4.85	0.02	439438	231284	163322	34647	6864	8353	5462	9404
N-ACETYL GLUTAMATE	3743.54	16365.63	4.37	2.13	0.03	29345	19945	9216	6955	3725	4278	2185	4787
BETA-ALANINE/ SARCOSINE	91671.91	74913.85	0.82	-0.29	0.06	80780	91893	63355	63628	108846	90706	88677	78459
ACETYL CARNITINE	39181822.43	50011154.73	1.28	0.35	0.02	51785865	46817629	57671187	43769938	36715835	46930381	34507449	38573625
BUTYRYL CARNITINE	10989190.86	22216590.40	2.02	1.02	0.04	23281029	34275288	20349287	10960758	6209266	15954557	10426047	11366893
CARNITINE	17380941.28	20010259.52	1.15	0.20	0.09	19353372	24317672	20443784	15926209	17614974	17707147	17482308	16719336
DECANOYL CARNITINE	490883.36	548368.46	1.12	0.16	0.30	409669	631700	617537	534568	216392	532370	603411	611361
DEOXY CARNITINE	9284299.12	12785918.10	1.38	0.46	0.05	12872972	16576548	13589899	8104253	9021978	9033336	9675447	9406436
GLUTARYL CARNITINE	981704.54	1432798.71	1.46	0.55	0.21	2356550	441786	2074583	858276	1281448	231013	1191874	1222484
HEXANOYL CARNITINE	552414.51	551980.84	1.00	0.00	0.50	536320	706744	563917	400942	293978	735529	652367	527783
ISOVALERYL CARNITINE	5247868.13	6564920.51	1.25	0.32	0.15	5476777	9096963	5894643	5791299	3309798	6506605	4751733	6423336
LAUROYL CARNITINE	3778620.13	3387789.32	0.90	-0.16	0.26	2913203	3313513	3251486	4072955	2862383	3663954	3326262	5261881
MYRISTOYL CARNITINE	15479121.16	13597060.52	0.88	-0.19	0.25	11212345	13101460	12478222	17596215	11917678	13733081	14373969	21891757
OCTANOYL CARNITINE	37207.04	48307.12	1.30	0.38	0.17	47228	64310	47805	33885	11812	42662	53838	40517
PALMITOYL CARNITINE	32044085.82	26802463.01	0.84	-0.26	0.06	24699726	27758386	24298110	30453629	26480154	32047564	30627373	39021252
STEAROYL CARNITINE	6525797.03	6020663.36	0.92	-0.12	0.00	5855506	6012325	6170557	6044265	6249662	6594223	6450664	6808639
HYDROXYPHENYL LACTATE	130803.42	108857.41	0.83	-0.26	0.33	63829	208000	87565	76035	96655	229692	100163	96704
BIOTIN	36542.92	60117.31	1.65	0.72	0.03	76055	71768	56206	36441	35566	47213	33956	29436

4-HYDROXY-3-METHOXYPHENYL GLYCOL	695.37	587.86	0.85	-0.24	0.15	620	785	430	517	627	858	595	701
CYSTINE	5659.06	7363.91	1.30	0.38	0.26	12520	10012	2357	4566	5118	6750	3535	7233
METHIONINE	42239.79	36129.59	0.86	-0.23	0.13	33983	46231	32960	31345	40835	52709	37737	37677
METHYLTHIO ADENOSINE	6056444.56	6733480.57	1.11	0.15	0.21	6800888	8638983	6183367	5310685	5332519	6060471	5759853	7072936
N-FORMYL METHIONINE	16584.27	34254.52	2.07	1.05	0.01	39693	44896	31087	21341	15552	18456	14368	17960
OPHTHALMATE	3709.99	8253.24	2.22	1.15	0.00	8395	9418	9050	6149	3685	3333	4139	3683
S-ADENOSYL HOMOCYSTEINE	12010.59	8799.02	0.73	-0.45	0.10	12207	10709	6673	5608	15166	14201	9417	9259
S-ADENOSYL METHIONINE	55222.81	55181.11	1.00	0.00	0.50	65423	62794	47462	45045	37135	61387	51906	70463
N-ACETYL ASPARAGINE	1684.75	2263.46	1.34	0.43	0.16	2307	3671	1942	1135	1500	1967	1429	1843
GUANOSINE DIPHOSPHATE MANNOSE	16593.29	18864.19	1.14	0.19	0.27	25797	22223	16272	11164	14851	14846	16651	20025
HEXOSE 1-PHOSPHATE	113215.54	245442.28	2.17	1.12	0.01	335430	244278	241999	160063	106752	117480	81595	147035
CDP-ETHANOLAMINE	24078.78	24033.75	1.00	0.00	0.49	28368	27581	21132	19055	19770	23557	23760	29229
CHOLINE	80043334.49	68587854.88	0.86	-0.22	0.01	74672020	68767490	64898536	66013373	88138276	79006551	79168544	73859966
ETHANOLAMINE	367715.85	268717.61	0.73	-0.45	0.00	273947	325051	210794	265079	363347	381603	345579	380334
GLYCEROL 3-PHOSPHATE	219753.97	125338.79	0.57	-0.81	0.01	102814	72992	112146	213403	220511	191623	217629	249252
CREATINE	96401666.88	95545159.86	0.99	-0.01	0.32	97089555	97602674	95284663	92203747	97581732	95923719	93132660	98968556
CYSTATHIONINE	117070.86	265703.10	2.27	1.18	0.19	680148	265170	68222	49273	26039	75433	102835	263976
GLYCINE	22801.02	42474.16	1.86	0.90	0.06	72570	44297	29993	23037	21858	21372	22376	25598
HYDROXY PYRUVATE	15432.33	18159.03	1.18	0.23	0.24	14006	11778	25568	21284	10827	18907	17842	14153
SERINE	375843.77	510662.61	1.36	0.44	0.06	673952	503484	501465	363750	344417	321094	357190	480675
THREONINE	74871.22	55277.42	0.74	-0.44	0.01	52978	42569	60775	64788	79418	69126	80716	70225
2-3 PHOSPHO GLYCERATE	109759.70	163989.44	1.49	0.58	0.06	193886	150926	180627	130519	37795	96509	151009	153725
LACTATE	75785.13	74568.52	0.98	-0.02	0.34	77460	74103	74547	72164	76852	74304	69969	82016
ACETYL-COA	3278.63	1786.46	0.54	-0.88	0.01	1739	1239	1367	2800	2610	3850	3790	2865
FRUCTOSE 6-PHOSPHATE	121082.45	253015.42	2.09	1.06	0.00	337782	246772	254500	173008	105834	122952	105345	150198
GLYCERALDEHYDE 3-PHOSPHATE	10793.08	7960.89	0.74	-0.44	0.13	6969	5543	6646	12685	6999	14256	12659	9258
PHOSPHOENOL PYRUVATE	38568.93	86765.71	2.25	1.17	0.03	136998	57631	84538	67896	11360	39321	52482	51112
4-IMIDAZOLE ACETATE	9850.88	11938.21	1.21	0.28	0.15	11019	16708	10401	9625	7774	11813	10069	9748
ANSERINE	127485.18	230611.64	1.81	0.86	0.05	266379	353629	173225	129214	100579	144263	112621	152478
CARNOSINE	4401213.08	6151158.24	1.40	0.48	0.20	7153331	10328237	3713049	3410016	2226825	5722756	2862347	6792924
HISTAMINE	116727.32	85590.82	0.73	-0.45	0.14	66615	86323	87753	101672	102443	186314	115685	62468
HISTIDINE	12778959.81	15679125.47	1.23	0.30	0.23	16690999	23882086	11033036	11110380	8629854	14571131	10313200	17601654
AMINOADIPATE	91533.36	77551.66	0.85	-0.24	0.25	98329	42217	91408	78252	104222	46381	112772	102760
PIPECOLATE	63931.66	341952.84	5.35	2.42	0.01	577283	304616	342270	143642	60383	38984	74027	82333
LYSINE	1803027.60	1402987.79	0.78	-0.36	0.05	1693746	1304389	1260782	1353034	1421635	1747481	1773031	2269964
N,N,N-TRIMETHYL LYSINE	16823604.08	11424182.32	0.68	-0.56	0.00	11514394	11121443	11374183	11686710	14736119	16953087	14989869	20615341
N-METHYL GLUTAMATE	91533.36	77551.66	0.85	-0.24	0.25	98329	42217	91408	78252	104222	46381	112772	102760
TRIMETHYLAMINE	2448233.97	2023243.24	0.83	-0.28	0.01	2212355	2012117	1879693	1988809	2797643	2441865	2348375	2205052
METHYLGUANIDINE	4796.77	6143.87	1.28	0.36	0.10	7993	4386	7371	4826	4404	4459	5756	4568
THYROTROPIN RELEASING HORMONE	7075.26	12084.53	1.71	0.77	0.03	15747	12766	12374	7452	9468	3074	7665	8093
NAD	1029099.17	916187.63	0.89	-0.17	0.31	1095088	1124416	709634	735613	643637	976931	937340	1558488
NADP	11315.53	8418.74	0.74	-0.43	0.09	7231	9507	8071	8866	5875	13468	13320	12600
NADPH	121634.38	116048.41	0.95	-0.07	0.36	136447	142112	95140	90494	110271	112612	130842	132812

NICOTINAMIDE	53132208.55	46599277.11	0.88	-0.19	0.02	45365337	46691177	42390753	51949842	57434455	51551741	50458856	53083783
NICOTINATE	791531.35	719657.37	0.91	-0.14	0.06	726271	741492	663376	747490	888545	727083	756823	793675
NICOTINIC ACID ADENINE DINUCLEOTIDE PHOSPHATE	21916.75	20561.80	0.94	-0.09	0.30	22721	24759	18545	16223	19124	19953	24024	24567
TRIGONELLINE	380212.79	623299.86	1.64	0.71	0.00	674723	587134	694535	536808	311322	291274	404577	513678
NADH	52222.90	22146.30	0.42	-1.24	0.00	19284	17957	19718	31626	72589	44548	51181	40573
AMP	1289181.48	1007651.96	0.78	-0.36	0.02	978358	1101948	852450	1097852	1037892	1342002	1342106	1434725
PANTOTHENATE	1510185.89	2409228.68	1.60	0.67	0.04	3001997	3097207	2149448	1388263	1454715	1838293	1342211	1405525
DEOXYRIBOSE	12373.86	11756.25	0.95	-0.07	0.32	9198	12151	11469	14207	12105	14495	11067	11828
GLUCONATE	68268.12	93900.00	1.38	0.46	0.07	122512	108536	87194	57358	64322	60208	71677	76865
GLYCERATE	169727.97	195051.93	1.15	0.20	0.16	228589	233197	162998	155424	165461	166147	150990	196314
RIBOSE 5-PHOSPHATE	74064.32	46523.69	0.63	-0.67	0.02	58810	38164	30886	58235	72413	83919	88647	51279
HIPPURATE	7401.61	10344.23	1.40	0.48	0.20	10720	5386	17027	8243	3956	5147	13559	6944
FRUCTOSE 1,6-BISPHOSPHATE	15462.65	49340.85	3.19	1.67	0.01	73423	55337	48477	20126	7979	17252	17938	18681
KYNURENINE	1533.08	1920.69	1.25	0.33	0.24	1407	3219	1780	1276	881	1766	1504	1982
N-ACETYLPHENYL ALANINE	24825.34	21660.91	0.87	-0.20	0.37	17826	38203	16388	14227	18343	47457	16171	17330
PHENYLALANINE	1266709.16	1238004.88	0.98	-0.03	0.47	1117766	1997075	974784	862395	1120560	1842666	1014161	1089450
TRYPTOPHAN	251188.28	223017.70	0.89	-0.17	0.34	197917	371314	175299	147540	221375	379702	196309	207367
TYROSINE	131575.16	121552.61	0.92	-0.11	0.43	92625	218987	90000	84598	92763	248534	91385	93620
O-PHOSPHO ETHANOL AMINE	52637.82	106066.02	2.02	1.01	0.02	144029	110502	121977	47756	52064	52519	55049	50919
GLYCOCHENO DEOXYCHOLATE	49621.59	59249.66	1.19	0.26	0.08	69304	56256	62563	48875	59112	52101	46498	40775
1-METHYL ADENOSINE	213681.28	420094.12	1.97	0.98	0.04	542069	608066	346667	183574	204176	194609	241588	214352
5'-DEOXY ADENOSINE	55582.13	51646.15	0.93	-0.11	0.23	55685	53986	53524	43390	66729	52188	47592	55819
ADENINE	2339463.26	1302758.45	0.56	-0.84	0.01	1689306	1161928	1134235	1225564	3035197	2213799	2323425	1785433
ADENOSINE	54069049.66	30173969.77	0.56	-0.84	0.01	39973288	28648531	25888629	26185431	68742644	50616086	56863714	40053754
ADENOSINE 3',5'-CYCLIC PHOSPHATE	14047.31	14215.13	1.01	0.02	0.45	12344	16969	12414	15133	11839	15513	14076	14761
ADENOSINE 3',5'-DIPHOSPHATE	530291.53	784852.53	1.48	0.57	0.02	834698	827313	826803	650597	291252	556424	589529	683961
ADENOSINE DIPHOSPHATE RIBOSE	404180.24	243499.16	0.60	-0.73	0.04	281379	211290	139805	341523	331723	399617	588883	296498
ALLANTOIN	15920.20	13620.29	0.86	-0.23	0.25	16140	6482	13517	18343	16763	10826	19736	16355
DEOXY GUANOSINE	1308.00	17258.87	13.19	3.72	0.01	18234	26657	22067	2077	1618	1574	1096	944
DGDP	365305.71	506907.91	1.39	0.47	0.02	532008	533048	518567	444009	229113	400751	388182	443176
GUANINE	17309.23	66460.29	3.84	1.94	0.02	75193	99314	73977	17357	17421	20442	14862	16511
GUANOSINE	120098.27	66334.74	0.55	-0.86	0.04	92026	79566	53115	40632	180282	100672	124604	74835
GUANOSINE DIPHOSPHATE	33285.23	32200.91	0.97	-0.05	0.38	36797	35998	29969	26040	26953	32377	36427	37383
GUANOSINE MONOPHOSPHATE	235741.41	192332.86	0.82	-0.29	0.01	179432	212925	189582	187392	201202	253113	233191	255460
HYPOXANTHINE	495978.90	347615.67	0.70	-0.51	0.00	366180	363661	288116	372506	541407	493027	519218	430263
INOSINE	4070734.79	3180721.47	0.78	-0.36	0.04	4050237	3093037	2701045	2878567	4946050	3580159	4009055	3747675
INOSINE MONOPHOSPHATE	159783.13	109502.17	0.69	-0.55	0.00	89486	127856	97471	123195	164070	175151	138046	161866
OXALATE	3665258.80	3643834.45	0.99	-0.01	0.44	3904985	3688660	3519053	3462640	3756899	3408675	3725607	3769854
URATE	31182.76	35598.73	1.14	0.19	0.26	37702	46435	34214	24043	26788	41751	34968	21224
XANTHINE	17096.24	14339.97	0.84	-0.25	0.10	17409	16075	13603	10273	20423	15506	16190	16266
XANTHOSINE-MONOPHOSPHATE	2907.07	1642.11	0.56	-0.82	0.00	2084	1905	1289	1290	3067	2931	3341	2290
CYTIDINE	37846.83	19864.91	0.52	-0.93	0.00	16118	19343	18081	25917	36452	34094	39862	40979
DEOXYURIDINE	19867.39	31843.61	1.60	0.68	0.02	37466	38872	30447	20589	16746	21910	20393	20420
THYMIDINE	1478.25	2751.86	1.86	0.90	0.03	3410	3352	3015	1231	1692	1695	1522	1005
URACIL	18641.27	12198.03	0.65	-0.61	0.01	12650	12928	11165	12049	24013	18167	17296	15089

UREIDOPROPIONATE	786.46	1765.34	2.24	1.17	0.01	2119	2196	1785	961	668	943	732	803
URIDINE	51076.52	28911.37	0.57	-0.82	0.00	27968	28691	26019	32967	53816	47012	52444	51034
URIDINE 5'-DIPHOSPHATE	23925.04	41919.21	1.75	0.81	0.09	64232	52880	34367	16198	15053	16598	27910	36140
URIDINE DIPHOSPHATE HEXOSE	2771010.15	3233394.17	1.17	0.22	0.17	3893516	3630399	3116601	2293061	2199116	2461013	3398853	3025058
URIDINE MONOPHOSPHATE	156306.81	163518.60	1.05	0.07	0.37	168229	146117	155237	184492	117676	150475	147331	209746
URIDINE TRIPHOSPHATE	1066.23	17716.99	16.62	4.05	0.01	28613	22073	16624	3558	0	1326	1352	1586
FAD	30704.49	22201.85	0.72	-0.47	0.01	24166	27848	16244	20550	28969	30861	28305	34683
HYPOTAURINE	1054.75	3745.35	3.55	1.83	0.04	6527	4753	2754	947	675	1821	767	956
TAURINE	1088928.95	1093569.38	1.00	0.01	0.48	1232780	1079862	1192234	869401	1146612	995155	1070941	1143008
CITRATE	23213583.84	24147004.87	1.04	0.06	0.31	26263727	27176237	22904138	20243918	21621713	22240977	23784181	25207464
ISOCITRATE	17951.04	15835.43	0.88	-0.18	0.12	18451	17868	13745	13278	19859	15314	18334	18297
MALATE	2366416.42	2274527.28	0.96	-0.06	0.22	2213051	2135220	2369731	2380107	2627258	2351513	2264930	2221964
SUCCINATE	159491.15	143628.13	0.90	-0.15	0.12	123084	131533	145133	174763	153701	173152	156601	154511
THIAMINE	150352.18	190794.75	1.27	0.34	0.19	155943	104408	286083	216745	188850	153327	102532	156700
SEROTONIN	206582.48	135302.02	0.65	-0.61	0.00	136252	145811	114117	145028	207162	223919	175352	219897
3-METHOXY TYRAMINE	275919.49	135330.73	0.49	-1.03	0.02	218820	135268	118190	69045	356449	316217	292864	138149
3,4-DIHYDROXY PHENYLGLYCOL	1425.12	1376.98	0.97	-0.05	0.43	1044	1543	930	1991	1371	1342	1513	1474
P-HYDROXY PHENYLACETATE	26489.20	15649.06	0.59	-0.76	0.01	18196	15731	14207	14462	34147	23795	27772	20242
TYRAMINE	7905.45	3588.23	0.45	-1.14	0.09	3520	5021	4150	1663	7113	15885	5532	3091
ARGININE	19126204.34	13510927.22	0.71	-0.50	0.00	13244609	13355716	11797723	15645661	17036342	18467446	20626990	20374038
CITRULLINE	33658.35	48368.63	1.44	0.52	0.11	78440	45151	41081	28802	25376	37573	34399	37285
ISOLEUCINE	113249.84	96815.48	0.85	-0.23	0.07	97008	115348	78684	96222	116819	118533	94454	123194
LEUCINE	267217.32	237290.93	0.89	-0.17	0.12	213586	279844	219629	236104	245516	260349	245413	317592
N-ACETYLLEUCINE	88750.88	134644.01	1.52	0.60	0.03	134872	158757	141837	103110	95686	113116	39049	107152
VALINE	276674.54	236511.08	0.85	-0.23	0.00	234916	246748	231338	233042	269206	284327	277059	276106
RIBOFLAVIN	23205.17	19100.94	0.82	-0.28	0.11	20124	17326	20405	18548	31473	23557	18529	19261
PYRIDOXAL	11824.20	9736.44	0.82	-0.28	0.04	8636	9512	9018	11780	11605	13137	12546	10008

Table 4.5. List of metabolites measured in PFSK-1 tumors and contralateral tissue using LC/MS

Metabolites	PFSK1 Contra Ave	PFSK-1 Tumor Ave	Fold Change	Log2 FC	T-test	PFSK-1 Tumor 1	PFSK-1 Tumor 2	PFSK-1 Tumor 3	PFSK-1 Tumor 4	PFSK-1 Contra 1	PFSK-1 Contra 2	PFSK-1 Contra 3	PFSK-1 Contra 4
(2-aminoethyl) phosphonate	10310.14	8857.78	0.86	-0.22	0.02	9362	8671	8658	8741	10182	11741	9359	9959
1-methyladenosine	5220.25	3050.25	0.58	-0.78	0.02	2074	3386	3783	2958	2905	5574	6344	6058
1-methylnicotinamide	19556.48	21453.61	1.10	0.13	0.18	17514	25726	22458	20117	17661	20100	19125	21340
1-methylnicotinamide	36162.94	357040.57	9.87	3.30	0.00	327878	566834	285652	247798	26033	34390	47070	37159
1,3-diaminopropane	2171.50	2644.00	1.22	0.28	0.07	2609	2404	2243	3320	1993	1964	2585	2144
10-hydroxydecanoate	160.57	130.75	0.81	-0.30	0.05	144	111	135	133	142	168	137	195
2-amino-2-methylpro-panoate	411246.01	309768.90	0.75	-0.41	0.02	311548	242402	317748	367377	348620	424518	475950	395896
2-hydroxybutyric acid	6475.79	8271.03	1.28	0.35	0.37	20173	3415	4617	4879	15648	2677	3856	3722
2-hydroxypyridine	38949.11	31102.88	0.80	-0.32	0.15	31798	24518	42273	25823	53333	40530	34737	27196
2-oxobutanoate	1058.31	1369.35	1.29	0.37	0.07	1297	1821	1108	1251	825	1040	1075	1292
2'-deoxycytidine 5'-monophosphate	347.35	410.98	1.18	0.24	0.19	429	364	399	452	303	217	354	516
2'-deoxyguanosine 5'-monophosphate	615267.22	419903.69	0.68	-0.55	0.03	453703	256445	363890	605577	581764	535520	707686	636099
2'-deoxyuridine 5' triphosphate	10716.90	5534.45	0.52	-0.95	0.02	5084	3215	4645	9193	14092	7162	8825	12789
3-(4-hydroxyphenyl) lactate	44727.65	48791.76	1.09	0.13	0.21	40170	49786	52438	52773	36387	40771	52814	48938
3-(4-hydroxyphenyl) pyruvate	456.76	333.83	0.73	-0.45	0.06	362	237	228	508	485	422	461	459
3-aminoisobutanoate	7070.00	6085.50	0.86	-0.22	0.10	6732	4425	6117	7068	6815	7069	8064	6332
3-hydroxy-3-methyl-glutarate	25662.77	28268.78	1.10	0.14	0.25	36045	26191	29642	21196	30772	26976	23173	21730
3-hydroxyanthranilate	2162.71	1984.69	0.92	-0.12	0.19	1766	2048	1953	2172	1918	2577	2282	1873
3-methoxytyramine	60975.00	34745.75	0.57	-0.81	0.07	30368	25297	47205	36113	17517	70163	71330	84890
3-methyglutaric acid	21436.07	18931.48	0.88	-0.18	0.15	17547	16702	20650	20826	26657	21813	18937	18338
3-methyl-2-oxindole	2739.15	3162.49	1.15	0.21	0.22	3120	2407	2735	4388	3444	2917	2500	2096
3,4-dihydroxy-l-phenylalanine	4316.75	8278.00	1.92	0.94	0.05	14267	6213	5716	6916	4695	4330	4632	3610
3,4-dihydroxyphenyl glycol	796.87	704.86	0.88	-0.18	0.31	569	513	750	988	1167	468	843	709
3,4-dihydroxypheny-lacetate	1572.39	2503.79	1.59	0.67	0.00	2422	2693	2312	2589	1544	1483	1415	1847
3',5'-cyclic amp	9467.75	6998.00	0.74	-0.44	0.10	8756	4180	6690	8366	6752	7831	12995	10293
4-aminobenzoate	1155247.97	3694018.44	3.20	1.68	0.04	525644	3716332	5595214	4938883	543762	1057643	1925099	1094488
4-aminobutanoate (GABA)	1781020.88	1445762.98	0.81	-0.30	0.03	1464177	1118035	1548311	1652530	1593468	1864470	1994189	1671957
4-guanidinobutanoate	1724346.25	1973120.25	1.14	0.19	0.10	2151353	1655114	1816063	2269951	1519767	1615933	1922226	1839459
4-imidazoleacetic acid	4590.16	3972.35	0.87	-0.21	0.03	4439	3313	3927	4210	4875	4500	4417	4568
4-pyridoxate	12255.75	25572.89	2.09	1.06	0.02	37787	22782	18610	23113	21932	7383	8806	10902
5-hydroxy indoleacetate	3933.65	4151.32	1.06	0.08	0.42	2281	4111	5545	4669	6052	3690	4011	1982
5'-deoxyadenosine	4664.05	5593.72	1.20	0.26	0.22	2818	7326	7269	4961	5468	4823	4207	4159
5'-methylthioadenosine	1831122.00	2134306.50	1.17	0.22	0.15	1632175	2353931	2357067	2194053	1364417	1795138	2374725	1790208
acetoin	1682.05	1131.64	0.67	-0.57	0.04	1474	609	1086	1358	1525	1302	2166	1735
acetylcarnitine	19251401.94	38696504.64	2.01	1.01	0.00	32326446	48516253	40482308	33461012	25946250	18766428	14178618	18114312
acetylcholine	171581.00	87335.25	0.51	-0.97	0.00	108223	44341	96541	100236	175686	122092	202816	185730
adenine	2514512.00	1584364.25	0.63	-0.67	0.03	2069123	962845	1322741	1982748	1801590	3141171	2578739	2536548
adenosine	36998857.02	23761029.27	0.64	-0.64	0.03	28837494	13072156	20754111	32380356	28453670	44978166	38006854	36556738
adenosine 2',3'-cyclic monophosphate	30597.26	22714.16	0.74	-0.43	0.08	26254	13371	23037	28195	25700	25150	41087	30451

adenosine 3',5'-cyclic monophosphate	12954.50	9892.00	0.76	-0.39	0.14	13492	5426	8119	12531	10195	10015	17357	14251
adenosine 3',5'-diphosphate	388784.15	387124.54	1.00	-0.01	0.49	399244	384231	403427	361597	331373	321115	541119	361530
adenosine 5'-diphosphoribose	112107.06	131068.73	1.17	0.23	0.26	184826	73073	149516	116859	143114	93219	126011	86084
adenosine 5'-monophosphate	560491.33	416305.06	0.74	-0.43	0.06	408176	250860	388712	617473	555558	492694	573932	619781
adenosine 5'-triphosphate	1208.26	5221.32	4.32	2.11	0.00	4080	6983	6290	3532	1118	794	1883	1038
agmatine	274634.24	843285.86	3.07	1.62	0.00	1006877	854792	820341	691133	392017	197590	296562	212367
alanine	2457499.21	2572198.43	1.05	0.07	0.08	2588725	2478400	2574620	2647049	2385844	2370596	2639557	2434000
alpha-aminoadipate	130601.99	241151.62	1.85	0.88	0.00	258950	174152	284160	247343	149954	108385	138445	125624
alpha-d-glucose	1206.86	2749.52	2.28	1.19	0.12	603	5918	3136	1341	617	1250	1507	1453
anserine	116232.05	75309.57	0.65	-0.63	0.03	99162	50199	61622	90256	88196	99075	147049	130609
anthranilate	6081.10	5292.89	0.87	-0.20	0.19	3983	4648	5309	7231	6442	5725	7049	5108
arginine	15749940.75	17001873.00	1.08	0.11	0.39	13275339	11170671	15875524	27685958	13656190	12487111	20482836	16373626
asparagine	22412.25	20127.00	0.90	-0.16	0.31	29067	13913	16284	21244	14262	25368	28025	21994
aspartate	1213595.12	1039588.37	0.86	-0.22	0.01	953061	1009271	1101275	1094746	1076413	1264893	1239531	1273543
azelaic acid	12944.15	15753.41	1.22	0.28	0.19	19926	10056	12988	20044	17491	11532	12197	10556
benzaldehyde	466645.25	475323.00	1.02	0.03	0.42	533218	377604	528473	461997	500545	440196	439014	486826
benzoate	28273.70	24600.28	0.87	-0.20	0.28	27988	17537	19659	33217	23550	17234	35721	36590
beta-alanine	46346.02	36820.55	0.79	-0.33	0.14	35427	29620	40426	41810	29769	46193	65794	43627
beta-nicotinamide adenine dinucleotide phosphate	14228.87	13769.00	0.97	-0.05	0.41	13670	11199	14800	15407	9262	13876	17165	16612
betaine	7487130.86	32289506.80	4.31	2.11	0.00	26397313	32330154	39373788	31056772	10870202	5802463	6924567	6351292
bilirubin	57040.25	904807.50	15.86	3.99	0.00	1340083	460900	1009317	808930	60999	51289	50610	65263
biotin	11916.18	43755.19	3.67	1.88	0.02	30546	78325	40655	25495	19758	10161	8423	9322
butyrylcarnitine	9563607.50	32021501.32	3.35	1.74	0.00	28310074	38480923	38120363	23174645	13609847	6803074	10774892	7066617
cadaverine	6937.00	19952.50	2.88	1.52	0.01	31945	13596	16493	17776	7889	5233	8891	5735
carnitine	21234503.31	30770326.20	1.45	0.54	0.00	30593260	26866046	32820404	32801595	23819319	18845697	22254889	20018109
cdp-ethanolamine	18850.32	14020.98	0.74	-0.43	0.05	14741	10999	13158	17185	14764	16631	24864	19142
choline	50420718.40	50689096.67	1.01	0.01	0.43	47581010	51369357	53748306	50057714	52209893	48718789	51692601	49061590
cis-4-hydroxy-proline	1951148.99	3620789.56	1.86	0.89	0.01	3416647	4676476	3836534	2553502	2024994	1906638	2196994	1675970
citicoline	578.44	706.25	1.22	0.29	0.21	705	906	564	650	394	514	969	437
citrate	5603314.33	6671730.53	1.19	0.25	0.16	8114409	5137662	5787244	7647607	6866558	3848777	6526201	5171722
citrate	3905.50	3865.50	0.99	-0.01	0.44	3984	4317	3651	3510	3955	4007	4231	3429
citrulline	34022.52	122456.39	3.60	1.85	0.00	110132	137057	133708	108929	33474	31401	33468	37747
cmp	11374.20	10587.37	0.93	-0.10	0.26	11374	8978	9388	12609	12018	9013	12603	11863
CMP	116045.77	85956.93	0.74	-0.43	0.10	84328	60263	73638	125600	93639	95257	159165	116121
creatine	80021091.29	72781440.13	0.91	-0.14	0.05	76875844	64066483	71586604	78596830	78688331	76205226	83698911	81491897
creatinine	5238483.77	3801864.38	0.73	-0.46	0.00	3276703	3857456	3710166	4363133	4569981	5527714	5594487	5261753
cystathionine	35784.95	114917.61	3.21	1.68	0.01	170315	75096	75157	139102	38400	16214	57118	31407
cytidine	608551.57	393234.03	0.65	-0.63	0.02	442041	211917	376795	542184	463613	652958	685661	631973
cytidine 5'-triphosphate	497.62	262.72	0.53	-0.92	0.00	367	184	231	268	479	503	472	537
cytosine	186276.30	496053.01	2.66	1.41	0.01	352249	769259	555220	307483	173868	151002	169172	251064
d-(-)-3-phosphoglyceric acid	44288.52	86418.66	1.95	0.96	0.04	140722	63812	70081	71061	32070	34635	70229	40220
d-glucosamine 6-phosphate	56927.38	60003.41	1.05	0.08	0.43	77617	41873	39540	80984	44563	34676	58501	89970
DAMP	50474.72	39294.79	0.78	-0.36	0.19	48610	26147	33109	49312	26743	42009	75668	57479
decanoylcarnitine	542729.60	1019438.78	1.88	0.91	0.04	1220431	821555	1020073	1015695	1147746	193773	482669	346730
dehydroascorbate	28242.45	30215.68	1.07	0.10	0.12	29388	28613	29678	33183	31427	27988	27252	26302
deoxycarnitine	8039111.01	9813492.29	1.22	0.29	0.04	10036340	7853832	9745082	11618715	7448761	7707023	9060559	7940101
deoxycytidine	121334.61	339807.08	2.80	1.49	0.01	256278	523245	375553	204152	114089	108168	111900	151181
deoxyribose	5381.00	5166.48	0.96	-0.06	0.23	4929	4700	5912	5126	5334	5375	5500	5314
deoxyuridine	5310.28	8861.80	1.67	0.74	0.00	8236	10156	9274	7781	4977	5108	5636	5520
diethanolamine	9602.39	6465.09	0.67	-0.57	0.03	7836	4560	4942	8522	10449	7676	8803	11481

dihydroorotate	4797.13	12274.61	2.56	1.36	0.00	12403	12765	13003	10927	5748	4397	4936	4108
dihydrouracil	5719.00	4140.75	0.72	-0.47	0.07	4129	2373	4307	5754	4889	4490	6876	6621
epinephrine	12342.95	25572.89	2.07	1.05	0.02	37787	22782	18610	23113	21932	7383	8806	11251
ethanolamine	547501.66	486165.97	0.89	-0.17	0.28	597074	301880	414386	631324	488914	403119	666292	631681
ethanolamine phosphate	74249.29	82813.83	1.12	0.16	0.13	92765	83900	85926	68663	59769	80623	77935	78670
ferulate	418405.00	388720.75	0.93	-0.11	0.32	518300	347508	380847	308228	449489	504459	397781	321891
flavin adenine dinucleotide	45537.25	35827.00	0.79	-0.35	0.10	41245	24011	31871	46181	34220	44750	56412	46767
fructose	15565.46	17834.88	1.15	0.20	0.03	17592	16742	17886	19119	17193	14201	16836	14031
fumarate	59812.63	62691.38	1.05	0.07	0.22	68956	61659	65943	54207	61103	58193	62438	57515
galactarate	841.58	2274.14	2.70	1.43	0.00	1755	3233	2229	1880	633	791	998	944
galactose 1-phosphate	478591.93	297336.86	0.62	-0.69	0.01	308467	188695	268416	423769	486447	437372	519013	471535
glucono-1,5-lactone	4600.14	4421.85	0.96	-0.06	0.34	4645	3434	4530	5079	4517	4086	5138	4660
gluconolactone	4600.14	4421.85	0.96	-0.06	0.34	4645	3434	4530	5079	4517	4086	5138	4660
glucose 1-phosphate	478760.86	294444.12	0.62	-0.70	0.00	301688	190643	272106	413340	487261	434164	520600	473018
glucose or fructose 6-phosphate	78673.11	61025.84	0.78	-0.37	0.05	65119	50099	53021	75864	82986	59482	92754	79470
glutamic acid	1220179.36	1251052.58	1.03	0.04	0.36	1145944	1426484	1294161	1137622	1143335	1156966	1280537	1299880
glutamic acid	17728254.32	14746156.03	0.83	-0.27	0.05	15473884	12531738	13722508	17256494	16033201	16316655	21035301	17527860
glutamine	25701059.61	22179708.48	0.86	-0.21	0.10	24575523	16882613	20951639	26309059	26705955	23656613	28892927	23548743
glutarylcarnitine	484376.31	684520.06	1.41	0.50	0.06	663939	488938	708203	877001	706679	367780	447538	415508
glutathione red	653647.74	692771.65	1.06	0.08	0.35	863180	487416	824869	595621	614050	647563	739023	613955
glycerate	2738.55	3345.65	1.22	0.29	0.11	3357	4101	2487	3438	1992	2531	3022	3409
glycerol	781.95	861.81	1.10	0.14	0.24	961	648	808	1030	800	595	846	888
glycerol 3-phosphate	203430.15	289289.37	1.42	0.51	0.04	217639	314974	280394	344151	147043	176984	287253	202441
glycine	17878.08	36197.83	2.02	1.02	0.00	42870	35012	32172	34738	20077	15503	19950	15983
glycochenodeoxycholate	199862.63	400447.80	2.00	1.00	0.08	699921	242235	269043	390592	419369	101272	135127	143681
glycolate	26127.15	30327.55	1.16	0.22	0.09	34494	31320	24302	31194	23826	23375	26471	30836
guanine	45775.25	50545.00	1.10	0.14	0.32	59195	37012	45290	60683	27866	37081	54052	64102
guanosine	98077.00	56930.50	0.58	-0.78	0.02	79199	34632	46822	67069	66070	127266	97737	101235
guanosine 5'-diphosphate	13082.09	10660.99	0.81	-0.30	0.12	11823	7534	9740	13548	12310	10785	17105	12128
guanosine 5'-diphospho-mannose	8331.01	6317.49	0.76	-0.40	0.08	5669	5144	6578	7879	5866	8176	11315	7967
guanosine 5'-monophosphate	184672.96	119367.95	0.65	-0.63	0.02	120460	74835	107060	175117	161992	171413	215063	190224
hexanoylcarnitine	1810593.75	5430932.89	3.00	1.58	0.00	6474783	4390746	5605898	5252305	2760126	941741	2014493	1526015
hippurate	1711.16	13033.64	7.62	2.93	0.05	510	24275	20824	6526	321	2725	2124	1675
histamine	48892.50	27262.00	0.56	-0.84	0.11	22933	27894	30420	27801	74685	27338	17391	76156
homocysteine	4192.00	3503.25	0.84	-0.26	0.31	5603	1918	1749	4743	2359	3002	5868	5539
homoserine	584330.58	1175611.24	2.01	1.01	0.00	1346063	943254	1112787	1300341	650004	477568	618521	591228
hydroxykynurenine	123.72	194.57	1.57	0.65	0.04	153	261	185	179	188	106	116	85
hydroxypyruvate	23669.28	59135.18	2.50	1.32	0.27	211469	5011	8100	11961	78050	3950	5886	6791
hypotaurine	3273.78	31976.36	9.77	3.29	0.01	16916	51699	38746	20544	5000	2078	2708	3309
hypoxanthine	700732.25	538750.50	0.77	-0.38	0.09	496559	392885	443087	822471	622303	651957	817754	710915
inosine	10857187.79	9921842.29	0.91	-0.13	0.15	8495976	9849319	9622040	11720034	11669679	10919044	11336005	9504023
inosine 5'-diphosphate	20556.86	23258.15	1.13	0.18	0.19	24592	20760	23511	24169	14982	18432	27688	21125
inosine 5'-monophosphate	84824.83	58506.53	0.69	-0.54	0.04	70140	36745	44568	82574	85421	68168	88625	97085
isocitric acid	5510.65	5490.07	1.00	-0.01	0.49	6206	3430	4846	7479	6835	4200	6278	4729
isoleucine	118871.90	208103.92	1.75	0.81	0.01	228829	250200	220445	132941	140588	103384	136702	94814
isovalerylcarnitine	8117677.87	24341242.70	3.00	1.58	0.00	17380947	24512605	34603573	20867845	11853042	5702825	9416012	5498833
ketoleucine	8289.26	8008.32	0.97	-0.05	0.43	9875	4848	7304	10007	7690	5474	9682	10310
kynurenic acid	912.79	4356.33	4.77	2.25	0.00	3090	5659	3179	5499	482	1020	592	1558
kynurenine	1587.04	3017.26	1.90	0.93	0.00	3433	3295	2522	2819	2215	1362	1455	1316
L-DOPA	4469.25	9734.00	2.18	1.12	0.05	17607	6893	7234	7202	4502	4335	4611	4429
lactate	33842.61	37198.28	1.10	0.14	0.01	37730	35263	37849	37952	35155	32165	35795	32256

lauroylcarnitine	3500054.43	4477028.15	1.28	0.36	0.27	6519389	3158268	3838880	4391575	7296892	1672141	2664326	2366859
leucine	209850.50	394880.50	1.88	0.91	0.01	383539	517231	435111	243641	209345	215546	228818	185693
lumichrome	3267.52	2777.43	0.85	-0.23	0.03	2796	2313	2960	3041	3030	3371	3635	3034
lumichrome	13620.15	13241.67	0.97	-0.04	0.45	14847	6467	17479	14174	11409	11862	18020	13190
malate	3648743.89	3826346.98	1.05	0.07	0.15	4190502	3586454	3781522	3746909	3881260	3597814	3667616	3448285
malonate	23673.45	59218.22	2.50	1.32	0.27	211469	5011	8432	11961	78050	3966	5886	6791
melatonin	7427.33	7404.28	1.00	0.00	0.49	8638	5145	6580	9254	7542	5023	9114	8030
meso-tartaric acid	105563.50	123920.00	1.17	0.23	0.05	128878	132313	123875	110614	92355	113006	124978	91915
methionine	157286.91	213996.11	1.36	0.44	0.07	160546	295123	194309	206006	106005	184761	160070	178311
methyl beta-D-galactoside	1060969.50	1087159.00	1.02	0.04	0.43	1440092	801419	1073972	1033153	1092000	1055115	1077491	1019272
methyl indole-3-acetate	7792.07	7728.44	0.99	-0.01	0.49	4450	4945	12374	9144	6116	5612	11076	8364
methylmalonate	64046.03	68931.61	1.08	0.11	0.28	54213	74405	75500	71609	50923	59781	79793	65687
n-acetyl-dl-glutamic acid	180165.06	170630.32	0.95	-0.08	0.26	192526	161325	161172	167499	206498	163057	194658	156448
N-acetyl-glutamic acid	7869.20	7738.29	0.98	-0.02	0.48	4450	4945	12374	9184	6132	5617	11221	8507
n-acetyl-l-aspartic acid	4793461.63	4127575.00	0.86	-0.22	0.04	4513625	3463844	3938655	4594175	4401913	5079862	4728745	4963326
N-acetylputrescine	10326.40	309931.77	30.01	4.91	0.00	299380	462425	325372	152550	15208	9272	9895	6930
N-acetylserine	461359.58	522044.32	1.13	0.18	0.19	401040	554842	622582	509713	578326	378116	460293	428703
N-alpha-acetyl-lysine	14146522.19	11014292.59	0.78	-0.36	0.09	14266904	7407136	7480351	14902780	14178155	13151740	15083106	14173088
N-amidino-aspartate (guanidinosuccinic acid)	34022.52	122456.39	3.60	1.85	0.00	110132	137057	133708	108929	33474	31401	33468	37747
N-methyl-aspartic acid	182065.56	159718.09	0.88	-0.19	0.11	173969	124706	146855	193342	183072	168392	199883	176915
N,N-dimethyl-arginine (sdma or adma)	1136337.48	2171259.11	1.91	0.93	0.00	2330673	2519208	1905313	1929842	912374	1157049	1209819	1266108
nad	73446.91	59734.60	0.81	-0.30	0.12	53600	44588	62759	77991	54761	68039	83885	87103
NAD	1158240.50	704657.00	0.61	-0.72	0.10	825463	403956	546441	1042768	468040	917510	1757664	1489748
nepsilon,nepsilon,nepsilon-trimethyllysine	14146522.19	11014292.59	0.78	-0.36	0.09	14266904	7407136	7480351	14902780	14178155	13151740	15083106	14173088
nicotinamide	23162293.50	17706011.75	0.76	-0.39	0.03	18155575	18000048	17509324	17159100	27593926	25187488	23391297	16476463
nicotinamide hypoxanthine dinucleotide	9345.86	7615.86	0.81	-0.30	0.09	6615	6781	7681	9387	9072	6823	11184	10305
nicotinamide mononucleotide	1902.72	1334.69	0.70	-0.51	0.00	1182	1428	1335	1393	1581	2216	1938	1876
nicotinate	368586.50	285102.50	0.77	-0.37	0.04	321800	254770	270468	293372	439877	355629	405215	273625
nicotinic acid adenine dinucleotide phosphate	1424.75	1204.46	0.85	-0.24	0.30	1005	718	1151	1943	996	1041	2269	1393
norleucine	1951148.99	3620789.56	1.86	0.89	0.01	3416647	4676476	3836534	2553502	2024994	1906638	2196994	1675970
normetanephrine	12342.95	25572.89	2.07	1.05	0.02	37787	22782	18610	23113	21932	7383	8806	11251
norvaline	253439.50	481777.00	1.90	0.93	0.00	441614	617340	498798	369356	259715	233209	246677	274157
o-succinyl-homoserine	17774.90	13678.18	0.77	-0.38	0.22	17163	6578	7213	23759	17284	10316	24237	19263
octanoylcarnitine	60040.72	190753.33	3.18	1.67	0.00	191350	184606	237422	149635	110203	21598	61539	46823
octopamine	6550.99	2356.49	0.36	-1.48	0.05	4105	1250	2779	1292	712	7376	7756	10360
ophthalmic acid	18208.68	48793.93	2.68	1.42	0.20	148532	14418	17108	15117	42289	9236	11903	9407
orotic acid	18336.85	45127.95	2.46	1.30	0.00	37560	56213	56024	30715	17004	19693	21644	15006
oxaloacetate	486.53	2709.09	5.57	2.48	0.00	3308	3583	2724	1221	768	332	462	384
oxoadipate	839.22	724.36	0.86	-0.21	0.26	535	1016	654	693	855	1152	500	850
palmitoylcarnitine	18138466.00	25161638.00	1.39	0.47	0.05	28485519	20443095	25516186	26201752	27462422	11893624	17176763	16021055
pantothenic acid	2197704.75	4757795.25	2.16	1.11	0.00	6358219	4689392	4126839	3856731	3375253	1779134	1670175	1966257
phenylalanine	839348.00	1192792.25	1.42	0.51	0.00	1097666	1422787	1227564	1023152	917864	829526	831353	778649
phenylethanolamine	291680.87	278289.09	0.95	-0.07	0.29	242300	251309	285322	334226	299716	266462	308559	291987
phospho(enol)pyruvic acid	20462.74	44174.32	2.16	1.11	0.03	73000	38828	37681	27189	17417	19320	31354	13759
phosphocreatine	157.16	1814.78	11.55	3.53	0.02	2116	3364	1318	461	192	52	234	150
phosphoribosyl pyrophosphate	1105.88	910.72	0.82	-0.28	0.25	949	832	706	1156	709	954	1841	920
proline	2181579.70	5189700.08	2.38	1.25	0.01	3731130	7384704	5946589	3696377	1912869	2293815	2466574	2053061
propionylcarnitine	15658616.47	35214408.03	2.25	1.17	0.00	35796112	31565260	37540965	35955296	19125951	13433319	17278603	12796593
pterin	2217.28	2655.13	1.20	0.26	0.06	2227	3112	2945	2337	2356	2070	2376	2067

putrescine	124729.83	541934.90	4.34	2.12	0.00	523421	542072	562222	540025	213049	79025	103979	102865
pyridoxamine	209588.52	212404.68	1.01	0.02	0.44	221266	184121	221125	223107	244342	204389	174764	214859
pyrimidine	58757.28	44802.28	0.76	-0.39	0.09	56333	36132	48575	38169	74605	63542	59015	37868
riboflavin	11058.54	11162.68	1.01	0.01	0.48	9525	9670	11357	14099	8639	13437	14341	7817
ribose 5-phosphate	32685.46	45763.27	1.40	0.49	0.04	57908	37913	51326	35906	38599	36408	23920	31815
s-(5'-adenosyl)-homo-cysteine	16535.00	12641.50	0.76	-0.39	0.06	13515	9187	12361	15503	13674	17356	14270	20840
s-(5'-adenosyl)-l-ho-mocysteine	53615.96	46742.42	0.87	-0.20	0.18	44692	34649	54105	53523	39729	54675	64605	55456
sarcosine	2457508.12	2560247.62	1.04	0.06	0.12	2597495	2425793	2572843	2644859	2390755	2362907	2638807	2437562
serine	365119.21	351812.64	0.96	-0.05	0.39	353366	267851	323606	462427	347952	327209	413104	372212
serotonin	104683.85	53952.53	0.52	-0.96	0.00	41510	54319	68721	51260	75732	110720	126859	105425
stearoylcarnitine	3810927.48	6038407.26	1.58	0.66	0.00	5816193	6668153	6456556	5212727	5091819	3038479	3790241	3323171
suberic acid	7494.68	6605.97	0.88	-0.18	0.15	8386	6566	5348	6124	6607	7277	7400	8695
taurine	1108070.20	1043734.16	0.94	-0.09	0.30	1080511	866428	1028548	1199449	874934	1103750	1333225	1120371
tetradecanoylcarnitine	11650349.50	13634152.50	1.17	0.23	0.29	19448022	10708709	12129428	12250451	19805802	6917550	10976500	8901546
thiamine	634045.25	1023039.75	1.61	0.69	0.06	565205	1501539	1170614	854801	415169	627077	799354	694581
thiamine pyrophos-phate	363.95	472.05	1.30	0.38	0.01	503	401	480	504	430	331	380	316
threonine	584330.58	1175611.24	2.01	1.01	0.00	1346063	943254	1112787	1300341	650004	477568	618521	591228
thymidine 5'-mono-phosphate	484.38	2167.95	4.48	2.16	0.00	2194	2773	2098	1608	431	377	611	518
thymine	3779.00	4217.50	1.12	0.16	0.16	3833	4204	4214	4619	4422	3096	3206	4392
thyrotropin releasing hormone	1547.00	556.00	0.36	-1.48	0.03	812	517	365	530	2389	2084	980	735
trans-4-hydroxy-proline	1231193.31	2201128.76	1.79	0.84	0.00	2352869	2599851	2208203	1643591	1471552	1108482	1334025	1010714
trans-aconitate	48024.39	51804.71	1.08	0.11	0.33	67756	39258	42939	57266	60302	39122	51597	41075
trigonelline	1244933.75	4211355.25	3.38	1.76	0.03	626278	4840581	6794279	4584283	560062	1294509	1902684	1222480
tryptophan	163905.18	244474.26	1.49	0.58	0.01	196383	271628	253782	256104	184793	116113	172432	182282
tyrosine	75114.12	101819.11	1.36	0.44	0.02	76415	108280	107785	114796	60368	73285	78202	88601
uracil	12161.75	12249.43	1.01	0.01	0.47	10807	10744	12611	14835	10720	11027	14399	12501
urate	107833.81	175154.80	1.62	0.70	0.00	152064	193683	182980	171893	110152	94177	109877	117129
uridine	70982.00	44387.75	0.63	-0.68	0.01	52117	28919	33868	62647	58005	75476	81060	69387
uridine 5'-diphosphate	55060.45	77495.75	1.41	0.49	0.11	107997	44020	67590	90375	41322	43218	78648	57053
uridine 5'-diphospho-glucuronic acid	6539.36	32944.77	5.04	2.33	0.00	35274	38456	34924	23125	5536	5697	8980	5945
uridine 5'-triphosphate	4489.57	2995.14	0.67	-0.58	0.03	2777	1814	2919	4471	3919	4043	5014	4982
uridine-5-monophos-phate	65577.88	74348.69	1.13	0.18	0.22	87256	63215	63977	82947	61609	45692	87310	67701
valine	488469.37	933616.65	1.91	0.93	0.00	866370	1128166	999216	740714	531929	447982	482617	491351
xanthine	8866.00	7538.25	0.85	-0.23	0.18	7596	6849	6078	9630	7386	11834	9146	7098
xanthosine	1176.28	1028.33	0.87	-0.19	0.29	909	610	988	1607	949	1097	1622	1038
xanthosine-monophos-phate	5203.85	3148.12	0.60	-0.73	0.00	3135	2070	3259	4128	4886	4947	5879	5104
xanthurenic acid	1343.37	2455.89	1.83	0.87	0.01	1820	3076	2266	2662	1090	1110	1459	1713

Table 4.6. List of metabolites measured in CNS-NPC tumors and contralateral tissue using ^{1}H NMR

Metabolite Name	Tumor							Contralateral						
	1950	1951	1952	2474	2475	2658	2659	1950	1951	1952	2474	2475	2658	2659
Alanine	1.66	1.13	1.72	1.71	1.91	0.83	0.54	1.46	1.98	1.89	1.69	1.71	1.28	1.13
Phospho-Choline	1.31	0.91	0.67	0.34	0.75	0.71	0.52	0.79	0.15	0.66	0.36	0.57	0.18	0.26
Creatine	5.04	4.85	4.22	4.02	4.29	4.80	3.53	4.57	4.97	4.88	4.01	4.57	4.31	4.16
Phospho-Creatine	5.20	2.82	3.90	3.45	3.14	4.27	4.24	4.47	3.09	3.70	3.44	3.44	4.31	4.19
Gamma aminobutyric acid	2.78	0.57	0.36	1.68	2.16	2.58	1.72	1.78	1.34	1.58	1.69	1.79	2.55	2.54
Glutamine	3.93	3.13	3.63	2.52	2.96	2.75	3.28	2.70	2.59	1.94	2.48	2.66	2.26	3.76
Glutamate	7.07	5.60	4.70	6.24	5.91	6.31	6.90	7.57	6.75	6.24	6.25	6.57	5.81	6.75
Glutathione	2.09	1.26	1.54	1.25	1.08	0.92	0.80	1.73	1.40	1.51	1.26	1.33	1.35	1.19
Glycine	1.61	2.72	1.58	0.62	1.78	1.17	1.05	1.25	0.97	0.74	0.69	1.13	0.99	1.00
Inositol	20.36	24.73	25.21	8.29	8.72	9.52	5.64	7.97	7.39	8.06	8.27	7.47	8.12	5.51
Lactate	5.24	4.37	3.98	5.75	6.13	4.13	4.44	5.41	5.87	5.98	5.74	6.61	4.50	5.56
N-acetyl aspartate	5.05	1.64	2.35	5.77	5.41	6.14	5.56	6.44	7.22	6.83	5.75	7.27	7.01	6.86
Taurine	15.89	10.80	11.85	12.89	13.10	8.33	7.24	13.82	14.83	13.99	12.87	13.98	8.70	10.90
Ascorbate (vitamin C)	4.73	5.00	4.72	0.99	0.76	1.58	0.67	1.61	1.29	1.94	1.00	1.07	2.49	1.40
N-acetyl-aspartyl-glutamate	0.75	1.16	1.98	0.56	0.36	0.42	0.60	0.00	0.40	0.31	0.59	0.05	0.26	0.29
Glycerophospho-choline	0.94	1.05	1.03	0.65	0.33	0.58	0.75	0.24	0.89	0.27	0.64	0.00	0.55	0.53
phosphoethanol-amine	6.26	8.49	8.70	3.21	3.11	2.61	1.74	3.71	3.40	3.65	3.16	3.27	2.05	2.20
Macromolecules	2.14	1.83	1.79	1.98	1.95	1.89	1.75	1.95	2.11	1.90	1.98	2.02	1.93	1.84
NAA + NAAG	5.80	2.80	4.33	6.33	5.77	6.56	6.15	6.44	7.62	7.13	6.33	7.32	7.27	7.15
Ratio of glu / glu + gln	0.64	0.64	0.56	0.71	0.67	0.70	0.68	0.74	0.72	0.76	0.72	0.71	0.72	0.64
Total choline	2.25	1.96	1.70	0.99	1.09	1.29	1.26	1.03	1.04	0.93	1.00	0.57	0.73	0.79
Total creatine	10.23	7.67	8.12	7.47	7.43	9.07	7.76	9.04	8.06	8.57	7.45	8.01	8.62	8.34

Table 4.7. List of metabolites measured in PFSK-1 tumors using ^{1}H NMR

Metabolite Name	Tumor									
	M1	M2	M3	M4	M8	M9	3295	3323	3322	3324
Alanine	2.012	3.364	3.556	0.385	2.5	0.236	0.773	1.273	0.951	0.217
Phospho-Choline	0.826	1.086	3.29	1.093	2.645	3.642	2.163	0.464	0.72	1.922
Creatine	4.635	3.568	4.248	2.14	3.71	3.646	4.943	4.216	4.159	3.761
Phospho-Creatine	4.239	2.84	1.43	1.811	2.299	9.17E-02	0.31	3.298	3.389	0
Gamma aminobutyric acid	2.208	0.901	8.82E-02	0	7.90E-02	0	0.489	1.727	1.283	0.21
Glutamine	2.98	2.403	0.198	0.808	1.35	0.318	1.474	3.11	2.44	1.601
Glutamate	7.361	6.29	7.95	2.644	6.49	4.446	6.262	6.067	7.544	4.308
Glutathione	1.481	1.563	1.004	1.969	0.969	1.714	1.698	0.963	1.398	0.958
Glycine	1.183	2.944	5.594	3.157	4.062	5.391	4.494	1.737	1.426	2.909
Inositol	9.605	4.524	3.899	4.482	6.541	1.616	5.287	5.362	5.134	5.342
Lactate	6.05	13.835	27.316	9.981	14.346	15.936	8.025	7.462	4.988	6.169
N-acetyl aspartate	6.355	3.616	1.164	1.348	1.391	0.792	1.745	5.159	5.676	1.595
Taurine	15.143	9.898	5.555	7.48	6.681	6.433	9.031	12.016	11.232	6.801
Ascorbate (vitamin C)	1.757	0.845	1.512	2.147	1.517	3.077	3.141	1.364	1.673	0.802
N-acetyl-aspartyl-glutamate	0.373	0	1.108	0.389	0.757	1.932	0.706	0.202	0.394	0.605
Glycerophospho-choline	0.48	0.229	0	0.508	0	0.876	0.769	0.325	0.219	0
phosphoethanol-amine	3.545	3.475	4.453	2.959	3.377	0.177	2.168	2.333	1.901	2.549
Macromolecules	1.987	2.029	2.472	1.464	2.049	1.941	1.671	1.924	1.854	1.599
NAA + NAAG	6.728	3.616	2.273	1.737	2.148	2.724	2.45	5.361	6.07	2.2
Ratio of glu / glu + gln	10.341	8.693	8.147	3.452	7.84	4.764	7.735	9.177	9.984	5.909
Total choline	1.306	1.315	3.29	1.6	2.645	4.519	2.932	0.789	0.94	1.922
Total creatine	8.874	6.408	5.679	3.951	6.009	3.737	5.254	7.514	7.547	3.761

Table 4.8. List of metabolites measured in PFSK-1 contralateral tissue using ^{1}H NMR

Metabolite Name	Contralateral									
	M1	M2	M3	M4	M8	M9	3295	3323	3322	3324
Alanine	1.674	7.50E-02	1.475	1.223	1.786	0.854	1.072	1.568	1.786	1.137
Phospho-Choline	0.95	0.611	0	0.568	0.381	0	0.273	0	0.42	0
Creatine	4.859	3.257	3.693	2.973	3.672	4.21	4.192	4.865	4.111	4.956
Phospho-Creatine	4.777	4.239	4.807	5.444	3.953	3.355	4.208	2.798	4.309	2.932
Gamma aminobutyric acid	2.613	1.714	2.86	1.356	1.592	0.752	2.858	2.755	1.189	1.728
Glutamine	2.78	2.655	2.481	2.738	2.556	3.298	3.384	3.294	3.221	3.522
Glutamate	8.272	9.573	7.651	8.409	8.036	4.813	7.406	6.686	9.113	7.443
Glutathione	1.612	1.635	1.036	1.433	1.118	0.645	1.386	1.224	1.076	1.345
Glycine	0.136	0.79	1.224	1.109	1.312	0	1.182	1.214	0.807	1.383
Inositol	8.779	6.852	6.86	6.616	7.587	5.368	5.545	5.237	4.878	7.527
Lactate	5.153	2.948	5.26	3.154	4.53	3.493	4.517	5.505	3.561	3.146
N-acetyl aspartate	8.302	8.235	7.192	7.852	7.497	4.822	6.709	6.862	7.632	7.143
Taurine	15.055	12.963	14.515	13.136	14.644	14.019	11.609	13.679	11.134	14.04
Ascorbate (vitamin C)	1.55	0.752	1.735	1.119	1.289	3.862	2.334	1.101	2.432	0
N-acetyl-aspartyl-glutamate	7.35E-03	0.569	0.421	0.102	0.442	0	0.311	0.365	0.429	0.712
Glycerophospho-choline	8.42E-02	0	1.001	5.46E-03	0	0.97	0.599	0.546	0.298	0.929
phosphoethanol-amine	2.379	2.555	2.635	2.863	4.854	0	1.664	2.598	2.058	1.321
Macromolecules	1.98	1.932	1.998	1.866	2.027	1.635	1.806	1.949	2.156	1.835
NAA + NAAG	8.31	8.804	7.613	7.954	7.938	4.822	7.02	7.226	8.061	7.855
Ratio of glu / glu + gln	11.052	12.229	10.133	11.147	10.591	8.111	10.79	9.98	12.335	10.965
Total choline	1.034	0.611	1.001	0.573	0.381	0.97	0.872	0.546	0.717	0.929
Total creatine	9.636	7.496	8.5	8.417	7.625	7.565	8.4	7.663	8.42	7.888

Table 4.9. List of metabolites measured in sham mice (left and right hemisphere) using ^{1}H NMR

Metabolite Name	Right Hemisphere					Left Hemisphere				
	3401	3402	3928	3957	3927	3401	3402	3928	3957	3927
Alanine	1.412	1.016	1.543	1.907	2.339	1.790	1.799	2.293	2.255	2.584
Phospho-Choline	0.294	0	0.451	0	0.535	0.000	8.53E-02	0	0.339	0.264
Creatine	3.835	3.671	3.715	2.999	4.07	4.709	4.351	3.883	3.549	3.35
Phospho-Creatine	4.957	3.939	4.808	4.359	3.4	3.830	3.556	3.372	3.547	3.395
Gamma aminobutyric acid	1.55	1.325	1.815	1.043	1.872	1.725	1.466	2.222	1.288	1.362
Glutamine	2.396	2.735	2.768	2.952	3.547	3.950	3.665	2.922	2.819	3.531
Glutamate	6.577	5.678	7.011	5.605	6.24	7.817	7.152	6.936	6.871	7.409
Glutathione	1.269	1.559	1.273	1.422	1.336	1.320	1.227	1.06	1.021	1.2
Glycine	2.412	0.805	0.695	1.021	1.686	1.882	1.305	1.634	1.415	1.513
Inositol	6.176	8.032	8.089	7.036	4.096	6.308	7.584	6.785	6.385	4.675
Lactate	5.06	4.832	5.076	5.692	5.54	7.046	6.887	6.441	5.494	5.26
N-acetyl aspartate	8.023	6.587	6.291	6.545	6.383	7.889	7.357	6.79	6.611	6.237
Taurine	15.887	13.886	13.518	13.569	14.638	15.891	14.014	13.037	13.03	13.368
Ascorbate (vitamin C)	3.206	1.815	1.219	1.706	0.231	1.802	1.04	0.378	1.224	0.806
N-acetyl-aspartyl-glutamate	0	0.627	0.244	0.427	0.43	0.408	0.743	0.192	0.444	0.244
Glycerophospho-choline	0	0.404	0	0.485	5.30E-05	0.480	0.164	0.388	0	0.342
phosphoethanol-amine	2.974	3.37	3.627	3.652	3.907	3.157	4.422	3.45	3.269	3.515
Macromolecules	2.258	1.994	1.771	2.016	1.912	2.148	2.162	1.877	1.985	1.919
NAA + NAAG	8.023	7.214	6.535	6.972	6.813	8.297	8.1	6.982	7.055	6.481
Ratio of glu / glu + gln	8.973	8.413	9.779	8.558	9.787	11.767	10.817	9.858	9.69	10.941
Total choline	0.294	0.404	0.451	0.485	0.535	0.480	0.249	0.388	0.339	0.605
Total creatine	8.792	7.61	8.523	7.358	7.47	8.539	7.908	7.255	7.095	6.745

Table 4.10. Lactate to Pyruvate ratio measured using ^{13}C pyruvate HMR

Sample Name	Tumor Lac/Pyr Ratio	Contralateral Lac/Pyr Ratio
EOR-2474 CNS-NPC	0.085955389	0.113691015
EOR-2475 CNS-NPC	0.126889249	0.143509972
EOR-2658 CNS-NPC	0.19061	0.256411478
EOR-2659 CNS-NPC	0.199927307	0.300660258
EOR-3323 PFSK-1	0.15814	0.12457
EOR-3322 PFSK-1	0.179813164	0.16611353
EOR-3324 PFSK-1	0.161851512	0.134261588

Bibliography

1. Aldiri, I. & Vetter, M. L. PRC2 during vertebrate organogenesis: A complex in transition. Developmental Biology 367, 91–99 (2012).

2. Alexandrov, L. B. et al. Mutational signatures associated with tobacco smoking in human cancer. Science 354, 618–622 (2016).

3. Andronesi, O. C. et al. Early changes in glioblastoma metabolism measured by MR spectroscopic imaging during combination of anti-angiogenic cediranib and chemoradiation therapy are associated with survival. npj Precision Oncology 1, (2017).

4. Ashton, T. M., McKenna, W. G., Kunz-Schughart, L. A. & Higgins, G. S. Oxidative Phosphorylation as an Emerging Target in Cancer Therapy. Clinical Cancer Research 24, 2482–2490 (2018).

5. Atchison, L., Ghias, A., Wilkinson, F., Bonini, N. & Atchison, M. L. Transcription factor YY1 functions as a PcG protein *in vivo*. EMBO J. 22, 1347–1358 (2003).

6. Babourina-Brooks, B. et al. Glycine: a non-invasive imaging biomarker to aid magnetic resonance spectroscopy in the prediction of survival in paediatric brain tumours. Oncotarget 9, (2018).

7. Barker, P. B. N-Acetyl Aspartate--A Neuronal Marker? Annals of Neurology 49, 423–424 (2001).

8. Bates, S. E. et al. Romidepsin in peripheral and cutaneous T-cell lymphoma: mechanistic implications from clinical and correlative data. British Journal of Haematology 170, 96–109 (2015).

9. Béguelin, W. et al. EZH2 enables germinal centre formation through epigenetic silencing of CDKN1A and an Rb-E2F1 feedback loop. Nature Communications 8, (2017).

10. Beloueche-Babari, M. et al. Identification of magnetic resonance detectable metabolic changes associated with inhibition of phosphoinositide 3-kinase signaling in human breast cancer cells. Mol. Cancer Ther. 5, 187–196 (2006).

11. Bodor, C. et al. EZH2 mutations are frequent and represent an early event in follicular lymphoma. Blood 122, 3165–3168 (2013).

12. Bomanji, J. B., Costa, D. C. & Ell, P. J. Clinical role of positron emission tomography in oncology. Lancet Oncol. 2, 157–164 (2001).

13. Bonev, B. & Cavalli, G. Organization and function of the 3D genome. Nature Reviews Genetics 17, 661 (2016).

14. Bonev, B. et al. Multiscale 3D Genome Rewiring during Mouse Neural Development. Cell 171, 557-572.e24 (2017).

15. Bonora, G., Plath, K. & Denholtz, M. A mechanistic link between gene regulation and genome architecture in mammalian development. Current Opinion in Genetics & Development 27, 92–101 (2014).

16. Boyer, L. A. et al. Polycomb complexes repress developmental regulators in murine embryonic stem cells. Nature 441, 349–353 (2006).

17. Bracken, A. P. Genome-wide mapping of Polycomb target genes unravels their roles in cell fate transitions. Genes & Development 20, 1123–1136 (2006).

18. Bradford, Y. M. et al. Zebrafish Models of Human Disease: Gaining Insight into Human Disease at ZFIN. ILAR Journal 58, 4–16 (2017).

19. Budanov, A. V. & Karin, M. p53 target genes sestrin1 and sestrin2 connect genotoxic stress and mTOR signaling. Cell 134, 451–460 (2008).

20. Burbee, D. G. et al. Epigenetic inactivation of RASSF1A in lung and breast cancers and malignant phenotype suppression. J. Natl. Cancer Inst. 93, 691–699 (2001).

21. Burger, P. C. Supratentorial Primitive Neuroectodermal Tumor (sPNET). Brain Pathology 16, 86–86 (2006).

22. Cairns, R. A., Harris, I., McCracken, S. & Mak, T. W. Cancer Cell Metabolism. Cold Spring Harbor Symposia on Quantitative Biology 76, 299–311 (2011).

23. Capper, D. et al. DNA methylation-based classification of central nervous system tumours. Nature 555, 469–474 (2018).

24. Caretti, G., Di Padova, M., Micales, B., Lyons, G. E. & Sartorelli, V. The Polycomb Ezh2 methyltransferase regulates muscle gene expression and skeletal muscle differentiation. Genes Dev. 18, 2627–2638 (2004).

25. Castillo, M., Smith, J. K. & Kwock, L. Correlation of Myo-inositol Levels and Grading of Cerebral Astrocytomas. Am. J. Neuroradiol. 21, 1645 (2000).

26. Chamberlain, S. J., Yee, D. & Magnuson, T. Polycomb Repressive Complex 2 Is Dispensable for Maintenance of Embryonic Stem Cell Pluripotency. Stem Cells 26, 1496–1505 (2008).

27. Chase, A. & Cross, N. C. P. Aberrations of EZH2 in Cancer. Clinical Cancer Research 17, 2613–2618 (2011).

28. Cheng, M., Bhujwalla, Z. M. & Glunde, K. Targeting Phospholipid Metabolism in Cancer. Front Oncol 6, 266 (2016).

29. Cheng, T. et al. Automated transfer and injection of hyperpolarized molecules with polarization measurement prior to *in vivo* NMR: Automated injection protocol for *in vivo* hyperpolarized NMR. NMR in Biomedicine 26, 1582–1588 (2013).

30. Choi, M. Y. & Kipps, T. J. Inhibitors of B-Cell Receptor Signaling for Patients With B-Cell Malignancies: The Cancer Journal 18, 404–410 (2012).

31. Coiffier, B. et al. Results from a pivotal, open-label, phase II study of romidepsin in relapsed or refractory peripheral T-cell lymphoma after prior systemic therapy. J. Clin. Oncol. 30, 631–636 (2012).

32. Comment, A. et al. Design and performance of a DNP prepolarizer coupled to a rodent MRI scanner. Concepts in Magnetic Resonance Part B: Magnetic Resonance Engineering 31B, 255–269 (2007).

33. Comment, A. & Merritt, M. E. Hyperpolarized Magnetic Resonance as a Sensitive Detector of Metabolic Function. Biochemistry 53, 7333–7357 (2014).

34. Conway, E., Healy, E. & Bracken, A. P. PRC2 mediated H3K27 methylations in cellular identity and cancer. Current Opinion in Cell Biology 37, 42–48 (2015).

35. Cordenonsi, M. et al. Links between tumor suppressors: p53 is required for TGF-beta gene responses by cooperating with Smads. Cell 113, 301–314 (2003).

36. Couronné, L., Bastard, C. & Bernard, O. A. TET2 and DNMT3A Mutations in Human T-Cell Lymphoma. New England Journal of Medicine 366, 95–96 (2012).

37. Cuperlovic-Culf, M., Ferguson, D., Culf, A., Morin, P. & Touaibia, M. 1H NMR Metabolomics Analysis of Glioblastoma Subtypes: CORRELATION BETWEEN METABOLOMICS AND GENE EXPRESSION CHARACTERISTICS. Journal of Biological Chemistry 287, 20164–20175 (2012).

38. Dammann, R. et al. Epigenetic inactivation of a RAS association domain family protein from the lung tumour suppressor locus 3p21.3. Nat. Genet. 25, 315–319 (2000).

39. Dansen, T. B. & Burgering, B. M. T. Unravelling the tumor-suppressive functions of FOXO proteins. Trends in Cell Biology 18, 421–429 (2008).

40. de Capoa, A. et al. DNA demethylation is directly related to tumour progression: evidence in normal, pre-malignant and malignant cells from uterine cervix samples. Oncol. Rep. 10, 545–549 (2003).

41. de Laat, W. & Duboule, D. Topology of mammalian developmental enhancers and their regulatory landscapes. Nature 502, 499–506 (2013).

42. Desandes, E., Guissou, S., Chastagner, P. & Lacour, B. Incidence and survival of children with central nervous system primitive tumors in the French National Registry of Childhood Solid Tumors. Neuro-Oncology 16, 975–983 (2014).

43. Dixon, J. R. et al. Chromatin architecture reorganization during stem cell differentiation. Nature 518, 331 (2015).

44. Donaldson-Collier, M. C. et al. EZH2 oncogenic mutations drive epigenetic, transcriptional, and structural changes within chromatin domains. Nature Genetics (2019). doi:10.1038/s41588-018-0338-y

45. Duncan, I. M. Polycomblike: a gene that appears to be required for the normal expression of the bithorax and antennapedia gene complexes of Drosophila melanogaster. Genetics 102, 49–70 (1982).

46. Duvic, M. et al. Phase 2 trial of oral vorinostat (suberoylanilide hydroxamic acid, SAHA) for refractory cutaneous T-cell lymphoma (CTCL). Blood 109, 31–39 (2007).

47. Ehrlich, M. DNA hypomethylation in cancer cells. Epigenomics 1, 239–259 (2009).

48. Engelman, J. A., Luo, J. & Cantley, L. C. The evolution of phosphatidylinositol 3-kinases as regulators of growth and metabolism. Nature Reviews Genetics 7, 606 (2006).

49. Esteller, M. et al. Inactivation of the DNA repair gene O6-methylguanine-DNA methyltransferase by promoter hypermethylation is associated with G to A mutations in K-ras in colorectal tumorigenesis. Cancer Res. 60, 2368–2371 (2000).

50. Feinberg, A. P. & Vogelstein, B. Hypomethylation distinguishes genes of some human cancers from their normal counterparts. Nature 301, 89–92 (1983).

51. Flavahan, W. A. et al. Insulator dysfunction and oncogene activation in IDH mutant gliomas. Nature 529, 110–114 (2016).

52. Fraser, J. et al. Hierarchical folding and reorganization of chromosomes are linked to transcriptional changes in cellular differentiation. Molecular Systems Biology 11, 852–852 (2015).

53. Fults, D., Pedone, C. A., Morse, H. G., Rose, J. W. & McKay, R. D. Establishment and characterization of a human primitive neuroectodermal tumor cell line from the cerebral hemisphere. J. Neuropathol. Exp. Neurol. 51, 272–280 (1992).

54. Gao, P. et al. Hypotaurine evokes a malignant phenotype in glioma through aberrant hypoxic signaling. Oncotarget 7, (2016).

55. Garcia-Bermudez, J. et al. Aspartate is a limiting metabolite for cancer cell proliferation under hypoxia and in tumours. Nature Cell Biology 20, 775–781 (2018).

56. Garcia-Manero, G., Stoltz, M. L., Ward, M. R., Kantarjian, H. & Sharma, S. A pilot pharmacokinetic study of oral azacitidine. Leukemia 22, 1680–1684 (2008).

57. Garcia-Manero, G. et al. Phase I Study of Oral Azacitidine in Myelodysplastic Syndromes, Chronic Myelomonocytic Leukemia, and Acute Myeloid Leukemia. Journal of Clinical Oncology 29, 2521–2527 (2011).

58. Gessi, M. et al. MYCN amplification predicts poor outcome for patients with supratentorial primitive neuroectodermal tumors of the central nervous system. Neuro-Oncology 16, 924–932 (2014).

59. Gil, J. & O'Loghlen, A. PRC1 complex diversity: where is it taking us? Trends in Cell Biology 24, 632–641 (2014).

60. Gnyszka, A., Jastrzebski, Z. & Flis, S. DNA methyltransferase inhibitors and their emerging role in epigenetic therapy of cancer. Anticancer Res. 33, 2989–2996 (2013).

61. Grau, D. J. et al. Compaction of chromatin by diverse Polycomb group proteins requires localized regions of high charge. Genes & Development 25, 2210–2221 (2011).

62. Gruetter, R. & Tkác, I. Field mapping without reference scan using asymmetric echo-planar techniques. Magn Reson Med 43, 319–323 (2000).

63. Gupta, R. K. et al. Relationships between choline magnetic resonance spectroscopy, apparent diffusion coefficient and quantitative histopathology in human glioma. J. Neurooncol. 50, 215–226 (2000).

64. Hanahan, D. & Weinberg, R. A. The Hallmarks of Cancer. Cell 100, 57–70 (2000).

65. Hanahan, D. & Weinberg, R. A. Hallmarks of Cancer: The Next Generation. Cell 144, 646–674 (2011).

66. Hattingen, E. et al. Myo-Inositol: a marker of reactive astrogliosis in glial tumors? NMR in Biomedicine 21, 233–241 (2008).

67. He, L. & Hannon, G. J. MicroRNAs: small RNAs with a big role in gene regulation. Nat. Rev. Genet. 5, 522–531 (2004).

68. Hegi, M. E. et al. MGMT Gene Silencing and Benefit from Temozolomide in Glioblastoma. New England Journal of Medicine 352, 997–1003 (2005).

69. Hennig, L. & Derkacheva, M. Diversity of Polycomb group complexes in plants: same rules, different players? Trends in Genetics 25, 414–423 (2009).

70. Herman, J. G. et al. Distinct patterns of inactivation of p15INK4B and p16INK4A characterize the major types of hematological malignancies. Cancer Res. 57, 837–841 (1997).

71. Hinz, S. et al. Expression profile of the polycomb group protein enhancer of Zeste homologue 2 and its prognostic relevance in renal cell carcinoma. J. Urol. 182, 2920–2925 (2009).

72. Hofmann, L., Slotboom, J., Boesch, C. & Kreis, R. Characterization of the macromolecule baseline in localized (1)H-MR spectra of human brain. Magn Reson Med 46, 855–863 (2001).

73. Horská, A. & Barker, P. B. Imaging of Brain Tumors: MR Spectroscopy and Metabolic Imaging. Neuroimaging Clinics of North America 20, 293–310 (2010).

74. Hříbková, H., Grabiec, M., Klemová, D., Slaninová, I. & Sun, Y.-M. Calcium signaling mediates five types of cell morphological changes to form neural rosettes. Journal of Cell Science 131, jcs206896 (2018).

75. Huang, Y., Chen, D.-H., Liu, B.-Y., Shen, W.-H. & Ruan, Y. Conservation and diversification of polycomb repressive complex 2 (PRC2) proteins in the green lineage. Briefings in Functional Genomics 16, 106–119 (2017).

76. Huidobro, C. et al. A DNA methylation signature associated with the epigenetic repression of glycine N-methyltransferase in human hepatocellular carcinoma. Journal of Molecular Medicine 91, 939–950 (2013).

77. Ingham, P. W. A gene that regulates the bithorax complex differentially in larval and adult cells of Drosophila. Cell 37, 815–823 (1984).

78. Italiano, A. et al. Tazemetostat, an EZH2 inhibitor, in relapsed or refractory B-cell non-Hodgkin lymphoma and advanced solid tumours: a first-in-human, open-label, phase 1 study. The Lancet Oncology 19, 649–659 (2018).

79. Jakacki, R. I. et al. Outcome and prognostic factors for children with supratentorial primitive neuroectodermal tumors treated with carboplatin during radiotherapy: A report from the Children's Oncology Group: Outcome for Children With Supratentorial PNET. Pediatric Blood & Cancer 62, 776–783 (2015).

80. Jansen, M. P. H. M. et al. Hallmarks of Aromatase Inhibitor Drug Resistance Revealed by Epigenetic Profiling in Breast Cancer. Cancer Research 73, 6632–6641 (2013).

81. Ji, L. et al. Mutant p53 Promotes Tumor Cell Malignancy by Both Positive and Negative Regulation of the Transforming Growth Factor β (TGF-β) Pathway. Journal of Biological Chemistry 290, 11729–11740 (2015).

82. Jones, P. A. & Baylin, S. B. The fundamental role of epigenetic events in cancer. Nat. Rev. Genet. 3, 415–428 (2002).

83. Kalo, E. et al. Mutant p53 Attenuates the SMAD-Dependent Transforming Growth Factor 1 (TGF-1) Signaling Pathway by Repressing the Expression of TGF- Receptor Type II. Molecular and Cellular Biology 27, 8228–8242 (2007).

84. Kanai, Y., Ushijima, S., Nakanishi, Y., Sakamoto, M. & Hirohashi, S. Mutation of the DNA methyltransferase (DNMT) 1 gene in human colorectal cancers. Cancer Lett. 192, 75–82 (2003).

85. Karube, K. et al. Integrating genomic alterations in diffuse large B-cell lymphoma identifies new relevant pathways and potential therapeutic targets. Leukemia 32, 675 (2017).

86. Kennison, J. A. The Polycomb and trithorax group proteins of Drosophila: trans-regulators of homeotic gene function. Annu. Rev. Genet. 29, 289–303 (1995).

87. Kim, H. et al. Serial Magnetic Resonance Spectroscopy Reveals a Direct Metabolic Effect of Cediranib in Glioblastoma. Cancer Research 71, 3745–3752 (2011).

88. King, A. A. et al. Long-term neurologic health and psychosocial function of adult survivors of childhood medulloblastoma/PNET: a report from the Childhood Cancer Survivor Study. Neuro-Oncology now242 (2016). doi:10.1093/neuonc/now242

89. Kahn, T. G. et al. Interdependence of PRC1 and PRC2 for recruitment to Polycomb Response Elements. Nucleic Acids Research gkw701 (2016). doi:10.1093/nar/gkw701

90. Kleer, C. G. et al. EZH2 is a marker of aggressive breast cancer and promotes neoplastic transformation of breast epithelial cells. Proc. Natl. Acad. Sci. U.S.A. 100, 11606–11611 (2003).

91. Krijger, P. H. L. & de Laat, W. Regulation of disease-associated gene expression in the 3D genome. Nature Reviews Molecular Cell Biology 17, 771 (2016).

92. Krijger, P. H. L. et al. Cell-of-Origin-Specific 3D Genome Structure Acquired during Somatic Cell Reprogramming. Cell Stem Cell 18, 597–610 (2016).

93. Kriks, S. et al. Dopamine neurons derived from human ES cells efficiently engraft in animal models of Parkinson's disease. Nature 480, 547–551 (2011).

94. Ku, M. et al. Genomewide Analysis of PRC1 and PRC2 Occupancy Identifies Two Classes of Bivalent Domains. PLoS Genetics 4, e1000242 (2008).

95. Kundu, S. et al. Polycomb Repressive Complex 1 Generates Discrete Compacted Domains that Change during Differentiation. Molecular Cell 65, 432-446.e5 (2017).

96. Landeira, D. & Fisher, A. G. Inactive yet indispensable: the tale of Jarid2. Trends in Cell Biology 21, 74–80 (2011).

97. Le Dily, F. et al. Distinct structural transitions of chromatin topological domains correlate with coordinated hormone-induced gene regulation. Genes & Development 28, 2151–2162 (2014).

98. Lee, S. Y. Temozolomide resistance in glioblastoma multiforme. Genes & Diseases 3, 198–210 (2016).

99. Lee, T. I. et al. Control of Developmental Regulators by Polycomb in Human Embryonic Stem Cells. Cell 125, 301–313 (2006).

100. Levine, S. S. et al. The Core of the Polycomb Repressive Complex Is Compositionally and Functionally Conserved in Flies and Humans. Molecular and Cellular Biology 22, 6070–6078 (2002).

101. Lewis, E. B. A gene complex controlling segmentation in Drosophila. Nature 276, 565 (1978).

102. Ley, T. J. et al. DNMT3A Mutations in Acute Myeloid Leukemia. New England Journal of Medicine 363, 2424–2433 (2010).

103. Li, M. H., Bouffet, E., Hawkins, C. E., Squire, J. A. & Huang, A. Molecular genetics of supratentorial primitive neuroectodermal tumors and pineoblastoma. Neurosurg Focus 19, E3 (2005).

104. Li, M. et al. Frequent Amplification of a chr19q13.41 MicroRNA Polycistron in Aggressive Primitive Neuroectodermal Brain Tumors. Cancer Cell 16, 533–546 (2009).

105. Lieberman-Aiden, E. et al. Comprehensive Mapping of Long-Range Interactions Reveals Folding Principles of the Human Genome. Science 326, 289–293 (2009).

106. Liu, P., Cheng, H., Roberts, T. M. & Zhao, J. J. Targeting the phosphoinositide 3-kinase pathway in cancer. Nature Reviews Drug Discovery 8, 627–644 (2009).

107. Lohr, J. G. et al. Discovery and prioritization of somatic mutations in diffuse large B-cell lymphoma (DLBCL) by whole-exome sequencing. Proceedings of the National Academy of Sciences 109, 3879–3884 (2012).

108. Louis, D. N. et al. The 2007 WHO Classification of Tumours of the Central Nervous System. Acta Neuropathologica 114, 97–109 (2007).

109. Mansour, M. R. et al. An oncogenic super-enhancer formed through somatic mutation of a noncoding intergenic element. Science 346, 1373–1377 (2014).

110. Marchesi, I. & Bagell, L. Role of Enhancer of Zeste Homolog 2 Polycomb Protein and Its Significance in Tumor Progression and Cell Differentiation. in Chromatin Remodelling (ed. Radzioch, D.) (InTech, 2013). doi:10.5772/55370

111. Margueron, R. et al. Ezh1 and Ezh2 Maintain Repressive Chromatin through Different Mechanisms. Molecular Cell 32, 503–518 (2008).

112. Margueron, R. & Reinberg, D. The Polycomb complex PRC2 and its mark in life. Nature 469, 343–349 (2011).

113. Maynard, J. et al. The use of 18F-fluorodeoxyglucose positron emission tomography (18F-FDG PET) as a pathway-specific biomarker with AZD8186, a PI3Kβ/δ inhibitor. EJNMMI Research 6, (2016).

114. McCabe, M. T. et al. EZH2 inhibition as a therapeutic strategy for lymphoma with EZH2-activating mutations. Nature 492, 108–112 (2012).

115. Mendenhall, E. M. et al. GC-Rich Sequence Elements Recruit PRC2 in Mammalian ES Cells. PLoS Genetics 6, e1001244 (2010).

116. Miki, Y. et al. A strong candidate for the breast and ovarian cancer susceptibility gene BRCA1. Science 266, 66–71 (1994).

117. Mlynárik, V., Gambarota, G., Frenkel, H. & Gruetter, R. Localized short-echo-time proton MR spectroscopy with full signal-intensity acquisition. Magnetic Resonance in Medicine 56, 965–970 (2006).

118. Modzelewska, K. et al. MEK Inhibitors Reverse Growth of Embryonal Brain Tumors Derived from Oligoneural Precursor Cells. Cell Reports 17, 1255–1264 (2016).

119. Morin, R. D. et al. Somatic mutations altering EZH2 (Tyr641) in follicular and diffuse large B-cell lymphomas of germinal-center origin. Nature Genetics 42, 181–185 (2010).

120. Musselman, C. A. et al. Molecular basis for H3K36me3 recognition by the Tudor domain of PHF1. Nature Structural & Molecular Biology 19, 1266–1272 (2012).

121. Nelson, S. J. et al. Association of early changes in 1 H MRSI parameters with survival for patients with newly diagnosed glioblastoma receiving a multimodality treatment regimen. Neuro-Oncology now159 (2016). doi:10.1093/neuonc/now159

122. Nora, E. P. et al. Spatial partitioning of the regulatory landscape of the X-inactivation centre. Nature 485, 381–385 (2012).

123. Northcott, P. A., Pfister, S. M. & Jones, D. T. W. Next-generation (epi)genetic drivers of childhood brain tumours and the outlook for targeted therapies. The Lancet Oncology 16, e293–e302 (2015).

124. Ntziachristos, P. et al. Genetic inactivation of the polycomb repressive complex 2 in T cell acute lymphoblastic leukemia. Nature Medicine 18, 298–302 (2012).

125. Okosun, J. et al. Integrated genomic analysis identifies recurrent mutations and evolution patterns driving the initiation and progression of follicular lymphoma. Nature Genetics 46, 176–181 (2014).

126. Olivero, W. C., Dulebohn, S. C. & Lister, J. R. The use of PET in evaluating patients with primary brain tumours: is it useful? Journal of Neurology, Neurosurgery & Psychiatry 58, 250–252 (1995).

127. Olsen, E. A. et al. Phase IIb multicenter trial of vorinostat in patients with persistent, progressive, or treatment refractory cutaneous T-cell lymphoma. J. Clin. Oncol. 25, 3109–3115 (2007).

128. Opstad, K. S., Bell, B. A., Griffiths, J. R. & Howe, F. A. An investigation of human brain tumour lipids by high-resolution magic angle spinning 1 H MRS and histological analysis. NMR in Biomedicine 21, 677–685 (2008).

129. Oricchio, E. et al. A Cell Engineering Strategy to Enhance the Safety of Stem Cell Therapies. Cell Reports 8, 1677–1685 (2014).

130. Padua, D. et al. TGFβ Primes Breast Tumors for Lung Metastasis Seeding through Angiopoietin-like 4. Cell 133, 66–77 (2008).

131. Pandey, R., Caflisch, L., Lodi, A., Brenner, A. J. & Tiziani, S. Metabolomic signature of brain cancer. Molecular Carcinogenesis 56, 2355–2371 (2017).

132. Park, J. Y. Promoter Hypermethylation in Prostate Cancer. Cancer Control 17, 245–255 (2010).

133. Peng, J. C. et al. Jarid2/Jumonji Coordinates Control of PRC2 Enzymatic Activity and Target Gene Occupancy in Pluripotent Cells. Cell 139, 1290–1302 (2009).

134. Peng, M., Yin, N. & Li, M. O. Sestrins Function as Guanine Nucleotide Dissociation Inhibitors for Rag GTPases to Control mTORC1 Signaling. Cell 159, 122–133 (2014).

135. Peters, G. J. et al. Basis for effective combination cancer chemotherapy with antimetabolites. Pharmacology & Therapeutics 87, 227–253 (2000).

136. Picard, D. et al. Markers of survival and metastatic potential in childhood CNS primitive neuro-ectodermal brain tumours: an integrative genomic analysis. The Lancet Oncology 13, 838–848 (2012).

137. Pirrotta, V. Polycombing the Genome: PcG, trxG, and Chromatin Silencing. Cell 93, 333–336 (1998).

138. Pomeroy, S. L. et al. Prediction of central nervous system embryonal tumour outcome based on gene expression. Nature 415, 436–442 (2002).

139. Provencher, S. W. Estimation of metabolite concentrations from localized *in vivo* proton NMR spectra. Magn Reson Med 30, 672–679 (1993).

140. Provencher, S. W. Estimation of metabolite concentrations from localized*in vivo* proton NMR spectra. Magnetic Resonance in Medicine 30, 672–679 (1993).

141. Qin, S. et al. Tudor domains of the PRC2 components PHF1 and PHF19 selectively bind to histone H3K36me3. Biochemical and Biophysical Research Communications 430, 547–553 (2013).

142. Quina, A. S., Buschbeck, M. & Di Croce, L. Chromatin structure and epigenetics. Biochemical Pharmacology 72, 1563–1569 (2006).

143. Reik, W. Stability and flexibility of epigenetic gene regulation in mammalian development. Nature 447, 425–432 (2007).

144. Reisman, D., Glaros, S. & Thompson, E. A. The SWI/SNF complex and cancer. Oncogene 28, 1653–1668 (2009).

145. Renault, V. M. et al. The pro-longevity gene FoxO3 is a direct target of the p53 tumor suppressor. Oncogene 30, 3207–3221 (2011).

146. Ricci, P. E. et al. Differentiating recurrent tumor from radiation necrosis: time for re-evaluation of positron emission tomography? AJNR Am J Neuroradiol 19, 407–413 (1998).

147. Richon, V. M. Cancer biology: mechanism of antitumour action of vorinostat (suberoylanilide hydroxamic acid), a novel histone deacetylase inhibitor. British Journal of Cancer 95, S2–S6 (2006).

148. Rutkowski, S. et al. Treatment of Early Childhood Medulloblastoma by Postoperative Chemotherapy Alone. New England Journal of Medicine 352, 978–986 (2005).

149. Santi, D. V., Norment, A. & Garrett, C. E. Covalent bond formation between a DNA-cytosine methyltransferase and DNA containing 5-azacytosine. Proc. Natl. Acad. Sci. U.S.A. 81, 6993–6997 (1984).

150. Sanulli, S. et al. Jarid2 Methylation via the PRC2 Complex Regulates H3K27me3 Deposition during Cell Differentiation. Molecular Cell 57, 769–783 (2015).

151. Schuettengruber, B. & Cavalli, G. Recruitment of Polycomb group complexes and their role in the dynamic regulation of cell fate choice. Development 136, 3531–3542 (2009).

152. Sexton, T. et al. Three-Dimensional Folding and Functional Organization Principles of the Drosophila Genome. Cell 148, 458–472 (2012).

153. Sharma, S., Kelly, T. K. & Jones, P. A. Epigenetics in cancer. Carcinogenesis 31, 27–36 (2010).

154. Shi, Y. et al. Structure of the PRC2 complex and application to drug discovery. Acta Pharmacologica Sinica 38, 963–976 (2017).

155. Shukla, V. & Lu, R. IRF4 and IRF8: governing the virtues of B lymphocytes. Frontiers in Biology 9, 269–282 (2014).

156. Siegel, George J. Basic neurochemistry: molecular, cellular, and medical aspects. (Lippincott Williams & Wilkins, 1999).

157. Simon, J. A. & Kingston, R. E. Mechanisms of Polycomb gene silencing: knowns and unknowns. Nature Reviews Molecular Cell Biology 10, 697–708 (2009).

158. Sing, A. et al. A Vertebrate Polycomb Response Element Governs Segmentation of the Posterior Hindbrain. Cell 138, 885–897 (2009).

159. Sneeringer, C. J. et al. Coordinated activities of wild-type plus mutant EZH2 drive tumor-associated hypertrimethylation of lysine 27 on histone H3 (H3K27) in human B-cell lymphomas. Proceedings of the National Academy of Sciences 107, 20980–20985 (2010).

160. Solaini, G., Sgarbi, G. & Baracca, A. Oxidative phosphorylation in cancer cells. Biochimica et Biophysica Acta (BBA) - Bioenergetics 1807, 534–542 (2011).

161. Souroullas, G. P. et al. An oncogenic Ezh2 mutation induces tumors through global redistribution of histone 3 lysine 27 trimethylation. Nature Medicine 22, 632–640 (2016).

162. Spence, T. et al. A novel C19MC amplified cell line links Lin28/let-7 to mTOR signaling in embryonal tumor with multilayered rosettes. Neuro-Oncology 16, 62–71 (2014).

163. Squazzo, S. L. Suz12 binds to silenced regions of the genomein a cell-type-specific manner. Genome Research 16, 890–900 (2006).

164. Steidl, E. et al. Myoinositol as a Biomarker in Recurrent Glioblastoma Treated with Bevacizumab: A 1H-Magnetic Resonance Spectroscopy Study. PLOS ONE 11, e0168113 (2016).

165. Stein, E. M. Molecular Pathways: IDH2 Mutations--Co-opting Cellular Metabolism for Malignant Transformation. Clinical Cancer Research 22, 16–19 (2016).

166. Stein, E. M. et al. Enasidenib in mutant IDH2 relapsed or refractory acute myeloid leukemia. Blood 130, 722–731 (2017).

167. Swigut, T. & Wysocka, J. H3K27 Demethylases, at Long Last. Cell 131, 29–32 (2007).

168. Takahashi, K. et al. Induction of Pluripotent Stem Cells from Adult Human Fibroblasts by Defined Factors. Cell 131, 861–872 (2007).

169. Taylor, M. D. et al. Radial glia cells are candidate stem cells of ependymoma. Cancer Cell 8, 323–335 (2005).

170. Timmermann, B. Role of Radiotherapy in the Treatment of Supratentorial Primitive Neuroectodermal Tumors in Childhood: Results of the Prospective German Brain Tumor Trials HIT 88/89 and 91. Journal of Clinical Oncology 20, 842–849 (2002).

171. Tirode, F. et al. Genomic Landscape of Ewing Sarcoma Defines an Aggressive Subtype with Co-Association of STAG2 and TP53 Mutations. Cancer Discovery 4, 1342–1353 (2014).

172. Ung, M., Ma, X., Johnson, K. C., Christensen, B. C. & Cheng, C. Effect of estrogen receptor α binding on functional DNA methylation in breast cancer. Epigenetics 9, 523–532 (2014).

173. Veneti, Z., Gkouskou, K. & Eliopoulos, A. Polycomb Repressor Complex 2 in Genomic Instability and Cancer. International Journal of Molecular Sciences 18, 1657 (2017).

174. Viale, A. et al. Oncogene ablation-resistant pancreatic cancer cells depend on mitochondrial function. Nature 514, 628–632 (2014).

175. Vivanco, I. & Sawyers, C. L. The phosphatidylinositol 3-Kinase–AKT pathway in human cancer. Nature Reviews Cancer 2, 489 (2002).

176. von Bueren, A. O. et al. Treatment of young children with localized medulloblastoma by chemotherapy alone: Results of the prospective, multicenter trial HIT 2000 confirming the prognostic impact of histology. Neuro-Oncology 13, 669–679 (2011).

177. Waddington, C. H. The epigenotype. 1942. Int J Epidemiol 41, 10–13 (2012).

178. Waitkus, M. S., Diplas, B. H. & Yan, H. Biological Role and Therapeutic Potential of IDH Mutations in Cancer. Cancer Cell 34, 186–195 (2018).

179. Wakabayashi, Y., Iwashima, A., Yamada, E. & Yamada, R. Enzymological evidence for the indispensability of small intestine in the synthesis of arginine from glutamate. II. N-acetylglutamate synthase. Arch. Biochem. Biophys. 291, 9–14 (1991).

180. Wakeling, L. A. et al. SIRT1 affects DNA methylation of polycomb group protein target genes, a hotspot of the epigenetic shift observed in ageing. Human Genomics 9, 14 (2015).

181. Walter, M. J. et al. Recurrent DNMT3A mutations in patients with myelodysplastic syndromes. Leukemia 25, 1153–1158 (2011).

182. Wang, J. et al. Imprinted X inactivation maintained by a mouse Polycomb group gene. Nature Genetics 28, 371–375 (2001).

183. Wang, L. et al. The Leukemogenicity of AML1-ETO Is Dependent on Site-Specific Lysine Acetylation. Science 333, 765–769 (2011).

184. Wang, L. et al. Hierarchical recruitment of polycomb group silencing complexes. Mol. Cell 14, 637–646 (2004).

185. Wang, X. et al. Targeting of Polycomb Repressive Complex 2 to RNA by Short Repeats of Consecutive Guanines. Molecular Cell 65, 1056-1067.e5 (2017).

186. Want, E. J. et al. Global metabolic profiling of animal and human tissues via UPLC-MS. Nature Protocols 8, 17–32 (2013).

187. Ward, C. S., Eriksson, P., Izquierdo-Garcia, J. L., Brandes, A. H. & Ronen, S. M. HDAC Inhibition Induces Increased Choline Uptake and Elevated Phosphocholine Levels in MCF7 Breast Cancer Cells. PLoS ONE 8, e62610 (2013).

188. Wenger, A. et al. Stem cell cultures derived from pediatric brain tumors accurately model the originating tumors. Oncotarget 8, (2017).

189. Whitcomb, S. J., Basu, A., Allis, C. D. & Bernstein, E. Polycomb Group proteins: an evolutionary perspective. Trends in Genetics 23, 494–502 (2007).

190. Whittaker, S. J. et al. Final Results From a Multicenter, International, Pivotal Study of Romidepsin in Refractory Cutaneous T-Cell Lymphoma. Journal of Clinical Oncology 28, 4485–4491 (2010).

191. Wilson, B. G. & Roberts, C. W. M. SWI/SNF nucleosome remodellers and cancer. Nat. Rev. Cancer 11, 481–492 (2011).

192. Yakushiji-Kaminatsui, N. et al. Similarities and differences in the regulation of HoxD genes during chick and mouse limb development. PLOS Biology 16, e3000004 (2018).

193. Yelamanchi, S. D. et al. A pathway map of glutamate metabolism. Journal of Cell Communication and Signaling 10, 69–75 (2016).

194. Yen, K. et al. AG-221, a First-in-Class Therapy Targeting Acute Myeloid Leukemia Harboring Oncogenic IDH2 Mutations. Cancer Discovery 7, 478–493 (2017).

195. Yin, J. et al. Potential Mechanisms Connecting Purine Metabolism and Cancer Therapy. Frontiers in Immunology 9, (2018).

196. Yoshihara, H. A. I. et al. High-field dissolution dynamic nuclear polarization of [1- 13 C]pyruvic acid. Phys. Chem. Chem. Phys. 18, 12409–12413 (2016).

197. Younes, A. et al. Combination of ibrutinib with rituximab, cyclophosphamide, doxorubicin, vincristine, and prednisone (R-CHOP) for treatment-naive patients with CD20-positive B-cell non-Hodgkin lymphoma: a non-randomised, phase 1b study. The Lancet Oncology 15, 1019–1026 (2014).

198. Yuan, M., Breitkopf, S. B., Yang, X. & Asara, J. M. A positive/negative ion–switching, targeted mass spectrometry–based metabolomics platform for bodily fluids, cells, and fresh and fixed tissue. Nature Protocols 7, 872–881 (2012).

199. Zhang, N. Role of methionine on epigenetic modification of DNA methylation and gene expression in animals. Animal Nutrition 4, 11–16 (2018).

200. Zhang, Z. et al. PRC2 Complexes with JARID2, MTF2, and esPRC2p48 in ES Cells to Modulate ES Cell Pluripotency and Somatic Cell Reprograming. STEM CELLS 29, 229–240 (2011).

201. Zinn, K. R. et al. Noninvasive bioluminescence imaging in small animals. ILAR J 49, 103–115 (2008).

202. Alabert, C. et al. Two distinct modes for propagation of histone PTMs across the cell cycle. Genes Dev. 29, 585–590 (2015).

203. Alizadeh, A. A. et al. Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. Nature 403, 503–511 (2000).

204. Casanova, M. et al. Polycomblike 2 facilitates the recruitment of PRC2 Polycomb group complexes to the inactive X chromosome and to target loci in embryonic stem cells. Development 138, 1471–1482 (2011).

205. Frick, M., Dörken, B. & Lenz, G. The molecular biology of diffuse large B-cell lymphoma. Ther Adv Hematol 2, 369–379 (2011).

206. Gröbner, S. N. et al. The landscape of genomic alterations across childhood cancers. Nature 555, 321–327 (2018).

207. Spivakov, M. & Fisher, A. G. Epigenetic signatures of stem-cell identity. Nat. Rev. Genet. 8, 263–271 (2007).

208. Sturm, D. et al. New Brain Tumor Entities Emerge from Molecular Classification of CNS-PNETs. Cell 164, 1060–1072 (2016).

209. Suzuki, M. M. & Bird, A. DNA methylation landscapes: provocative insights from epigenomics. Nat. Rev. Genet. 9, 465–476 (2008).

210. Gilmore, T. D., Kalaitzidis, D., Liang, M.-C. & Starczynowski, D. T. The c-Rel transcription factor and B-cell proliferation: a deal with the devil. Oncogene 23, 2275–2286 (2004).

211. Henry, P.-G. et al. Proton-observed carbon-edited NMR spectroscopy in strongly coupled second-order spin systems. Magn Reson Med 55, 250–257 (2006).

212. Bellomo, C., Caja, L. & Moustakas, A. Transforming growth factor β as regulator of cancer stemness and metastasis. Br. J. Cancer 115, 761–769 (2016).

213. Thon, N., Kreth, S. & Kreth, F.-W. Personalized treatment strategies in glioblastoma: MGMT promoter methylation status. Onco Targets Ther 6, 1363–1372 (2013).

214. Vander Heiden, M. G., Cantley, L. C. & Thompson, C. B. Understanding the Warburg effect: the metabolic requirements of cell proliferation. Science 324, 1029–1033 (2009).

215. Warburg, O. On the origin of cancer cells. Science 123, 309–314 (1956).

216. Warburg, O., Wind, F. & Negelein, E. THE METABOLISM OF TUMORS IN THE BODY. J. Gen. Physiol. 8, 519–530 (1927).

217. Weischenfeldt, J. et al. Pan-cancer analysis of somatic copy-number alterations implicates IRS4 and IGF2 in enhancer hijacking. Nat. Genet. 49, 65–74 (2017).

218. Wright, G. et al. A gene expression-based method to diagnose clinically distinct subgroups of diffuse large B cell lymphoma. Proc. Natl. Acad. Sci. U.S.A. 100, 9991–9996 (2003).

219. Yap, D. B. et al. Somatic mutations at EZH2 Y641 act dominantly through a mechanism of selectively altered PRC2 catalytic activity, to increase H3K27 trimethylation. Blood 117, 2451–2459 (2011).

220. Prasad, P., Lennartsson, A. & Ekwall, K. The roles of SNF2/SWI2 nucleosome remodeling enzymes in blood cell differentiation and leukemia. Biomed Res Int 2015, 347571 (2015).

221. Rolfe, D. F. & Brown, G. C. Cellular energy utilization and molecular origin of standard metabolic rate in mammals. Physiol. Rev. 77, 731–758 (1997).

222. Zhang, Z. et al. PRC2 Complexes with JARID2, MTF2, and esPRC2p48 in ES Cells to Modulate ES Cell Pluripotency and Somatic Cell Reprograming. STEM CELLS 29, 229–240 (2011).

223. Molecular biology of the cell. (Garland Science, 2002).

224. Shen, H. & Laird, P. W. Interplay between the cancer genome and epigenome. Cell 153, 38–55 (2013).

Appendix 1

EZH2 Oncogenic Mutations Drive Epigenetic, Transcriptional, and Structural Changes within Chromatin Domains

Personal Contribution Statement

I contributed to the design of experiments, data analysis, and interpretation of results presented in the manuscript.

Figure 1:

- I designed, optimized experiments and generated Hi-C libraries of all new data used for all analyses of chromatin interactions published throughout the manuscript
- I performed western blot experiments and immunoblot acquisition

Figure 2

- I performed ChIP-sequencing experiments for H3K27me3
- I extracted mRNA from cell lines for RNA-sequencing

Figure 3

- I extracted mRNA from cell lines for qPCR validation of selected TADs

Figure 5

- I extracted mRNA from cell lines for qPCR validation and RNA sequencing of lymphoma cell lines treated with EZH2-inhibitors
- I performed western blot experiments and immunoblot acquisition

Figure 6

- I designed, optimized experiments and generated UMI-4C libraries
- I optimized and performed DNA-fluorescent in situ hybridization (FISH) targeting TAD6.139
- I assisted in image acquisition of STORM super-resolution imaging

nature genetics

ARTICLES

https://doi.org/10.1038/s41588-018-0338-y

EZH2 oncogenic mutations drive epigenetic, transcriptional, and structural changes within chromatin domains

Maria C. Donaldson-Collier[1,5], Stephanie Sungalee[1,5], Marie Zufferey[2,3,5], Daniele Tavernari[2,3,5], Natalya Katanayeva[1], Elena Battistello[1,2,3], Marco Mina[2,3], Kyle M. Douglass[4], Timo Rey[4], Franck Raynaud[2,3], Suliana Manley[4], Giovanni Ciriello[2,3*] and Elisa Oricchio[1*]

Chromatin is organized into topologically associating domains (TADs) enriched in distinct histone marks. In cancer, gain-of-function mutations in the gene encoding the enhancer of zeste homolog 2 protein (EZH2) lead to a genome-wide increase in histone-3 Lys27 trimethylation (H3K27me3) associated with transcriptional repression. However, the effects of these epigenetic changes on the structure and function of chromatin domains have not been explored. Here, we found a functional interplay between TADs and epigenetic and transcriptional changes mediated by mutated *EZH2*. Altered EZH2 (p.Tyr646* (EZH2^{Y646X})) led to silencing of entire domains, synergistically inactivating multiple tumor suppressors. Intra-TAD gene silencing was coupled with changes of interactions between gene promoter regions. Notably, gene expression and chromatin interactions were restored by pharmacological inhibition of EZH2^{Y646X}. Our results indicate that EZH2^{Y646X} alters the topology and function of chromatin domains to promote synergistic oncogenic programs.

In eukaryotic cells, chromatin progressively folds into a hierarchical structure including loops made of relatively short sequences (~1–10 kb), medium-sized domains (100–1,000 kb), also known as topologically associating domains (TADs), and large chromatin compartments (>1 Mb)[1–3]. In particular, TADs are highly conserved among species and cell types[4]. They are preferentially decorated by either active (H3K36me3) or inactive (H3K27me3) epigenetic marks, which can affect the transcriptional activity within the domains[5–7], and have thus been proposed as functional units regulating gene expression in development and differentiation[8,9]. Interestingly, cancer cells can hijack these chromatin domains to drive oncogenic transformation, either through chromosomal aberrations[10,11] or as a consequence of selected mutations affecting the cancer epigenome[12,13].

Mutations of epigenetic modifiers have been detected in several tumor types[14], and alterations in these proteins vary the global epigenetic status of tumor cells. However, their effects on chromatin organization remain unclear. Among recurrent alterations of chromatin remodeling factors, *EZH2* gain-of-function mutations affecting Tyr646 (RefSeq isoform: NM_004456) are common in non-Hodgkin lymphoma[15] and other tumor types[16,17]. EZH2 is the histone lysine methyltransferase regulating H3K27 methylation, and it is part of the Polycomb repressor complex 2 (ref. [18]). EZH2^{Y646X} aberrantly increases H3K27me3, promoting transcriptional repression[19,20]. This process has been implicated in stalling B-cell differentiation[21] and epigenetic silencing of tumor suppressors[22]. Given the associations among H3K27me3, chromatin subcompartments, and TADs[23], we wondered whether oncogenic *EZH2* mutations might lead to chromatin structural and/or functional modifications. Here, we show that epigenetic and transcriptional changes induced by *EZH2* mutations are more concordant within TADs than expected and modulate intra-TAD interactions proximal to gene promoters. Intra-TAD structural changes are associated with transcriptional inactivation of entire domains. Our results illustrate how cancer-associated epigenetic alterations can act beyond single genes and modify the activity of entire chromatin domains.

Results

Hi-C chromatin maps of EZH2^{WT} and EZH2^{Y646X} lymphomas. EZH2 p.Tyr646* gain-of-function alterations lead to a genome-wide increase in H3K27me3 (Supplementary Fig. 1a). To establish whether this global accumulation of H3K27me3 modifies the genome topology on a similarly broad scale, we performed high-throughput chromatin conformation capture (Hi-C) in two lymphoma cell lines (Karpas422 and WSU-DLCL2) expressing the mutant form of EZH2 (EZH2^{Y646X}) and a lymphoma cell line (OCI-Ly19) expressing the wild-type EZH2 (EZH2^{WT}) protein (Supplementary Table 1 and Supplementary Note). Contact maps of EZH2^{Y646X} and EZH2^{WT} cells were binned in regions of 50 kb and compared with multiple metrics (Fig. 1). For each pair of maps, we compared the overall distribution of chromosomal contacts by using the stratum-adjusted correlation coefficient (SCC)[24] (Fig. 1a); the fraction of 1-Mb regions assigned to the same compartment (A or B)[3] (Fig. 1b); the similarity among TADs[25–27] (Fig. 1c); and the fraction of bin interactions that were significant in both maps, also known as the cell interactome[28,29] (Fig. 1d). To build a reference scale of values for each metric, we compared Hi-C maps of EZH2^{Y646X} lymphoma cell lines with Hi-C maps of endothelial cells (HUVEC), fetal fibroblasts (IMR90), and normal lymphoblastoid cells (GM12878). Moreover, we used randomized contact maps or

[1]Swiss Institute for Experimental Cancer Research (ISREC), School of Life Science, École Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland. [2]Department of Computational Biology, University of Lausanne (UNIL), Lausanne, Switzerland. [3]Swiss Institute of Bioinformatics (SIB), Lausanne, Switzerland. [4]Institute of Physics, École Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland. [5]These authors contributed equally: Maria C. Donaldson-Collier, Stephanie Sungalee, Marie Zufferey, Daniele Tavernari. *e-mail: giovanni.ciriello@unil.ch; elisa.oricchio@epfl.ch

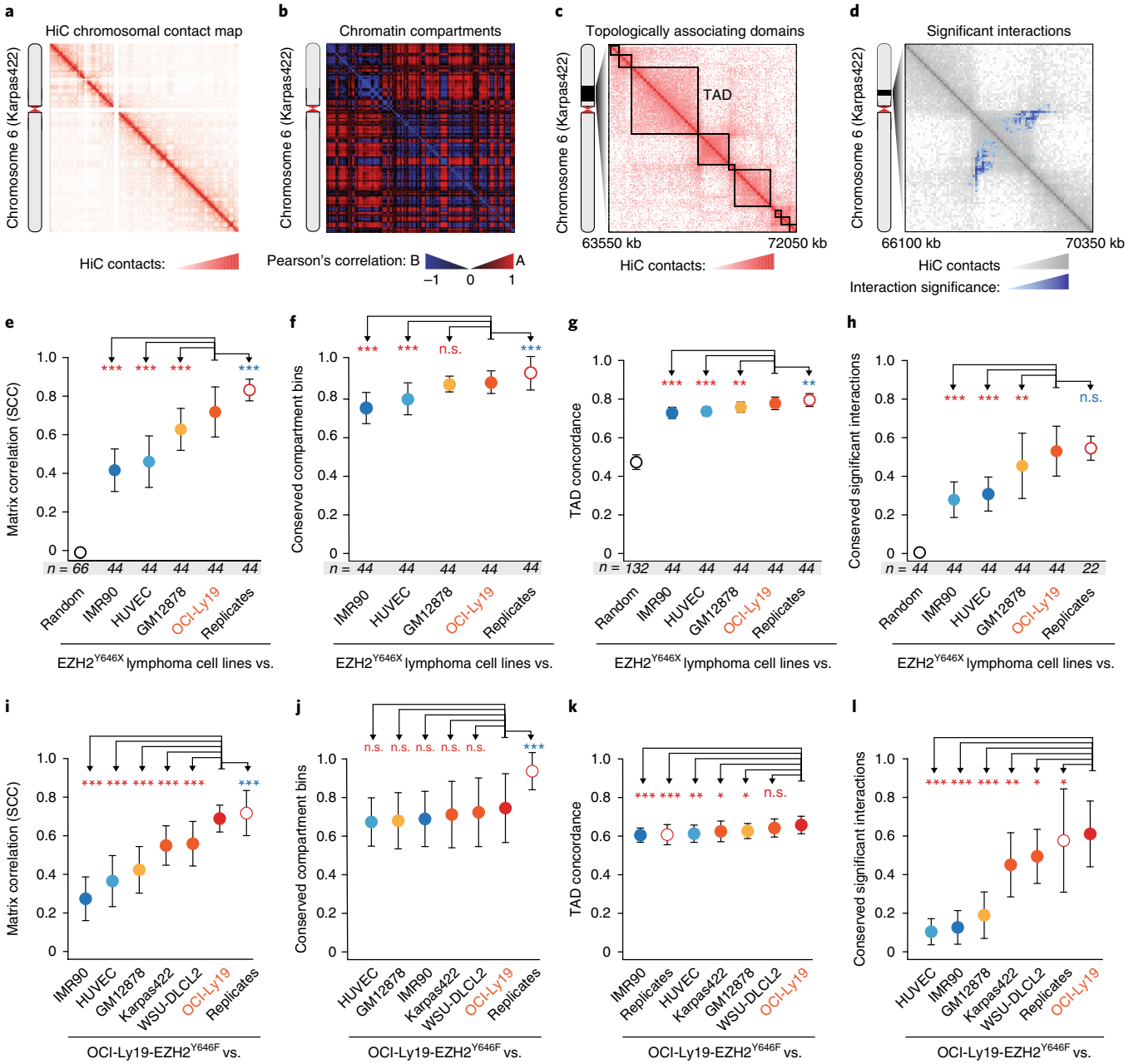


Fig. 1 | Comparison of the genome 3D organization in EZH2^{WT} and EZH2^{Y646X} cells. a–d, Chromatin structural elements identifiable by Hi-C; representative images for Chr.6 of the GM12878 cell line: intrachromosomal contact map (**a**); correlation of observed/expected contact ratios for each pair of 1-Mb bins, with the plaid (blue/red) pattern showing chromatin compartments A or B (**b**); TADs (**c**); significant interactions determined by using HiC-DC, with all pairs of 50-kb bins within a 2-Mb window tested (**d**). **e–h**, Comparison between the intrachromosomal contact maps of EZH2^{Y646X} cell lines (Karpas422 and WSU-DLCL2) and of the indicated cell lines, based on the overall Hi-C matrix correlation using the SCC (**e**); fraction of 1-Mb bins assigned to the same compartment (**f**); TAD similarity measured by measure of concordance (MoC) (**g**); and fraction of shared significant interactions between 50-kb bins determined by HiC-DC (**h**). **i–l**, Comparison between intrachromosomal contact maps of OCI-Ly19-EZH2^{Y646F} and of the indicated cell lines as described above, that is, based on SCC (**i**), compartment conservation (**j**), TAD similarity (**k**), and the fraction of shared significant interactions (**l**). For all comparisons, the number of compared intrachromosomal maps is indicated below the graphs (*n*), when not indicated $n = 22$; the dotd represent the mean values, and error bars are ±1 standard deviation. *P* values were computed with two-tailed Wilcoxon test (Supplementary Table 1). $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$; n.s., not significant.

TADs and independent experimental replicates of EZH2^{Y646X} lymphoma cell lines to provide the lowest and highest expected values for each metric, respectively. The similarity ranks obtained with all metrics were highly concordant and invariably showed that the contact maps from EZH2^{Y646X} lymphoma were on average more similar to those from EZH2^{WT} cell lines than to any other cell model that we tested, except for their replicates (Fig. 1e–h and Supplementary Fig. 1b–e).

To validate this observation and allay nonspecific effects due to heterogeneity among samples, we generated a syngeneic lymphoma cell line (OCI-Ly19) expressing the altered protein EZH2^{Y646F} (c.1937A>T, p.Tyr646Phe). We confirmed that expression of

EZH2^{Y646F} increased H3K27me3 across the genome, as compared with both the parental cell line and OCI-Ly19 transduced with an empty vector, whereas expression of EZH2 remained similar in the three conditions (Supplementary Fig. 1f,g). Hi-C contact maps derived for OCI-Ly19-EZH2^{Y646F} and OCI-Ly19 cells were as similar to or more similar than maps derived from independent replicates of the same cell line (Fig. 1i–l and Supplementary Fig. 1h). We conclude that although a certain diversity exists for a limited number of specific contacts and contact domains, on average, the genome topology of *EZH2*-mutated and wild-type lymphoma cells is highly similar.

EZH2^{Y646X} inactivates selected TADs. Given the observed consistency between the chromatin structures of EZH2^{Y646X} and EZH2^{WT} cell lines, we asked whether an association might exist between chromatin structural elements and the epigenetic and transcriptional changes induced by *EZH2* gain-of-function mutations. In particular, we focused on chromatin subcompartments and TADs, given that both have been associated with H3K27me3 (ref. [23]).

To assess changes in H3K27me3 specifically induced by *EZH2* mutations, we analyzed H3K27me3 levels in OCI-Ly19 cells expressing either EZH2^{Y646F} or EZH2^{WT}. The distribution of H3K27me3 in chromatin subcompartments confirmed the previously reported enrichment in H3K27me3 within the B1 compartment[23] (Supplementary Fig. 2a). However, in OCI-Ly19-EZH2^{Y646F}, we observed a significant but similar increase in H3K27me3 across all subcompartments (Supplementary Fig. 2a,b), thus suggesting that EZH2^{Y646X} does not induce epigenetic changes specifically within one subcompartment.

Next, we tested whether differences in H3K27me3 and gene expression were more concordant within TADs than expected and whether this association might lead to the activation or inactivation of specific domains (Fig. 2a). First, we extracted a consensus list of 2,038 TADs that comprised at least three expressed genes and were conserved among three EZH2^{Y646X} cell lines (Karpas422, WSU-DLCL2, and OCI-Ly19-EZH2^{Y646F}) and one EZH2^{WT} cell line (OCI-Ly19) (Supplementary Table 2). H3K27me3 chromatin immunoprecipitation and high-throughput sequencing (ChIP-seq) data were compared both between OCI-Ly19-EZH2^{WT} and OCI-Ly19-EZH2^{Y646F} and between OCI-Ly19-EZH2^{WT} and two cell lines (Karpas422 and WSU-DLCL2) with endogenous *EZH2* mutations. The TAD mean H3K27me3 values in OCI-Ly19-EZH2^{Y646F} were highly correlated with those observed in Karpas422 and WSU-DLCL2 (Supplementary Fig. 2c); thus, we decided to focus on the OCI-Ly19-EZH2^{WT} and OCI-Ly19-EZH2^{Y646F} ChIP-seq datasets, because they reflected H3K27me3 changes directly associated with the *EZH2* mutation. The distribution of TAD mean H3K27me3 values exhibited a positive shift in EZH2^{Y646X} compared with EZH2^{WT} cells (Fig. 2b). H3K27me3 profiles of loci within the same TAD were more correlated than profiles of loci belonging to distinct adjacent domains in both cell lines (Supplementary Fig. 2d,e). Moreover, H3K27me3 fold changes between OCI-Ly19 mutated cells and wild-type cells were also more correlated within the same TAD than between adjacent TADs (Fig. 2c), and this trend was independent of the distance between the genomic loci (Fig. 2d). Overall, these results indicate that the EZH2-mediated increase in H3K27me3 is associated with the compartmentalization of the chromatin in TADs.

Interestingly, the distribution of H3K27me3 fold changes (EZH2^{Y646X} versus EZH2^{WT}) within TADs exhibited a greater number of high fold changes than random distributions obtained by permuting 50-kb or 100-kb intervals of H3K27me3 ChIP-seq reads (Fig. 2e). This trend was independent of the subcompartment to which each TAD was assigned (Supplementary Fig. 2f) and of the interval size used in the permutations (Supplementary Fig. 2g). Moreover, by comparing matched TAD mean H3K27me3 values in EZH2^{WT} and EZH2^{Y646F} OCI-Ly19 cells, we found that H3K27me3 did not proportionally increase in all domains, but fold changes were greater in domains that had low histone methylation in EZH2^{WT} cells than in domains that were already enriched in H3K27me3 (Fig. 2f). Therefore, the increase in H3K27me3 mediated by EZH2 alterations within specific domains was not randomly distributed but were dependent on the initial level of H3K27me3 (Fig. 2g).

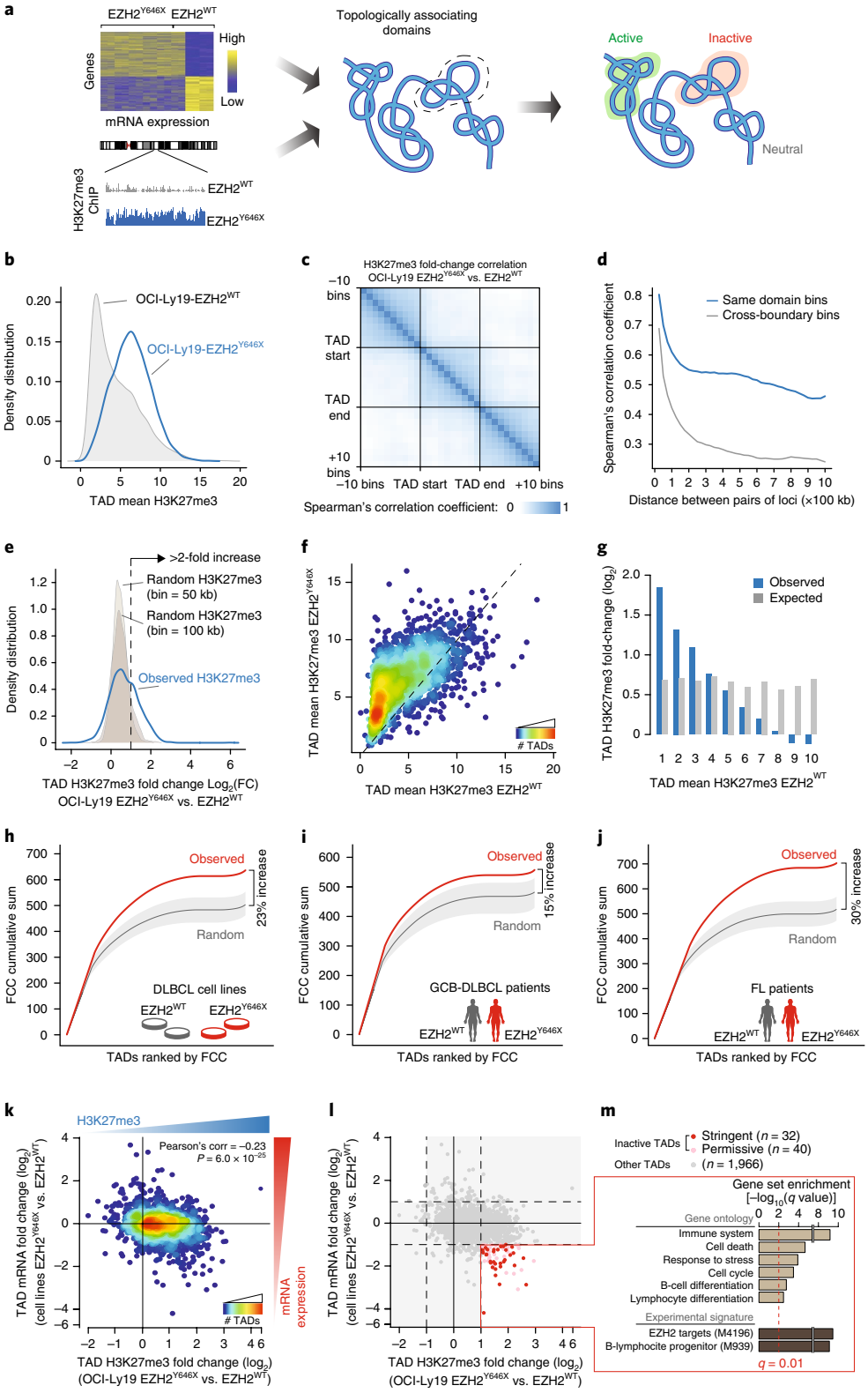
Next, we assessed the effects of H3K27me3 changes on transcription, in OCI-Ly19-EZH2^{Y646F} and OCI-Ly19-EZH2^{WT} cell lines. In particular, we wondered whether the observed correlation between H3K27me3 changes within TADs might be reflected in concordant mRNA expression changes in genes within the same TAD. For this purpose, we scored each TAD on the basis of the number and magnitude of concordant gene fold changes within the domain (mRNA fold-change concordance (FCC) scores; Supplementary Fig. 3a). For this analysis, we selected only TADs for which at least three genes had detectable mRNA expression in either OCI-Ly19-EZH2^{Y646F} or OCI-Ly19-EZH2^{WT}. The observed FCC scores were compared with those obtained after randomly permuting gene-to-TAD assignments. The observed FCC scores were higher than expected on the basis of random permutations, thus indicating that changes in gene expression in OCI-Ly19-EZH2^{Y646F} compared with OCI-Ly19-EZH2^{WT} were significantly concordant within TADs (Supplementary Fig. 3b).

To validate this finding in larger cohorts of B-cell lymphoma samples, we collected independent mRNA expression data for diffuse large B-cell lymphoma (DLBCL) cell lines[30] including Karpas422 and WSU-DLCL2 (EZH2^{WT} $n=2$, EZH2^{Y646X} $n=6$; each sample was analyzed in duplicate, GSE40792), and primary samples from germinal center (GC) patients with DLBCL[21] (EZH2^{WT} $n=30$,

Fig. 2 | Epigenetic and transcriptional changes in EZH2^{Y646X} cells occur within TADs. a, Study design: H3K27me3 ChIP-seq, mRNA expression, and Hi-C data in EZH2^{Y646X} and EZH2^{WT} lymphomas were integrated to identify epigenetically and transcriptionally inactive, neutral, and active TADs. **b**, Distribution of mean H3K27me3 signal within TADs in OCI-Ly19-EZH2^{WT} and OCI-Ly19-EZH2^{Y646F} cells. **c**, Correlation across $n=2,038$ TADs of H3K27me3 fold changes in loci within the same TAD and loci separated by one TAD boundary. **d**, Correlation (*y* axis) of H3K27me3 fold changes between all 25-kb bin pairs within the same TADs and bin pairs separated by one TAD boundary as a function of their distance (*x* axis). **e**, Distributions of mean H3K27me3 fold changes for $n=2,038$ TADs and of fold changes obtained by permuting H3K27me3 bins (50 kb and 100 kb). **f**, Correlation between TAD mean H3K27me3 ($n=2,038$ TADs) in OCI-Ly19-EZH2^{WT} (*x* axis) and OCI-Ly19-EZH2^{Y646F} (*y* axis). **g**, TAD mean H3K27me3 fold changes ($n=2,038$ TADs) between OCI-Ly19-EZH2^{Y646F} and OCI-Ly19-EZH2^{WT}; H3K27me3 values in OCI-Ly19-EZH2^{WT} were binned ($bin_i=(i-1,i)$), and TAD mean H3K27me3 fold changes in each bin are reported. Expected values were determined by TAD permutation. **h**–**j**, Cumulative sum curves of mRNA FCC for EZH2^{Y646X} ($n=12$) versus EZH2^{WT} cells ($n=4$) in 2,038 TADs (**h**), GC-DLBCL EZH2^{Y646X} ($n=7$) versus EZH2^{WT} ($n=30$) patient samples in 2,038 TADs (**i**), and FL EZH2^{Y646X} ($n=6$) versus EZH2^{WT} ($n=17$) patient samples in 1,908 TADs (**j**). Observed curves are compared with random curves obtained after 10,000 gene-to-TAD assignment permutations (gray area delimits minimum and maximum values; dark line is the mean). **k**,**l**, TAD mean H3K27me3 fold changes between OCI-Ly19-EZH2^{Y646F} and OCI-Ly19-EZH2^{WT} cells (*x* axis) versus the TAD mRNA-expression fold changes in EZH2^{Y646X} versus EZH2^{WT} cell lines (*y* axis) for $n=2,038$ TADs; in **l**, inactive TADs ($\log_2(FC_{H3K27me3})>1$ and $\log_2(FC_{mRNA})<-1$) are highlighted. **m**, Gene-set enrichment analysis for $n=283$ genes within inactive TADs according to Gene Ontology categories ($n=5,337$) and experimentally derived gene sets ($n=3,409$). Gene-set enrichment was tested by hypergeometric test (one-sided) and adjusted by false discovery rate.

EZH2^{Y646X} $n=7$, GSE23501) and patients with follicular lymphoma (FL)[31] (EZH2^{WT} $n=17$, EZH2^{Y646X} $n=6$, PRJNA278311). Across all datasets, we verified that mRNA expression was more correlated between genes in the same TAD than between genes in different TADs, independently of their genomic distance (Supplementary Fig. 3c–e). Furthermore, the FCC scores obtained by comparing EZH2^{Y646X} and EZH2^{WT} cases were always greater than random, with an overall increase ranging between 15% and 33% (increase in area under the curve; Fig. 2h–j and Supplementary Fig. 3b). Across all the expression datasets that we analyzed, expression changes associated with *EZH2* mutations were more concordant within TADs than expected.

Next, we compared matched mean H3K27me3 and mRNA expression changes within TADs in OCI-Ly19-EZH2^{Y646F} and



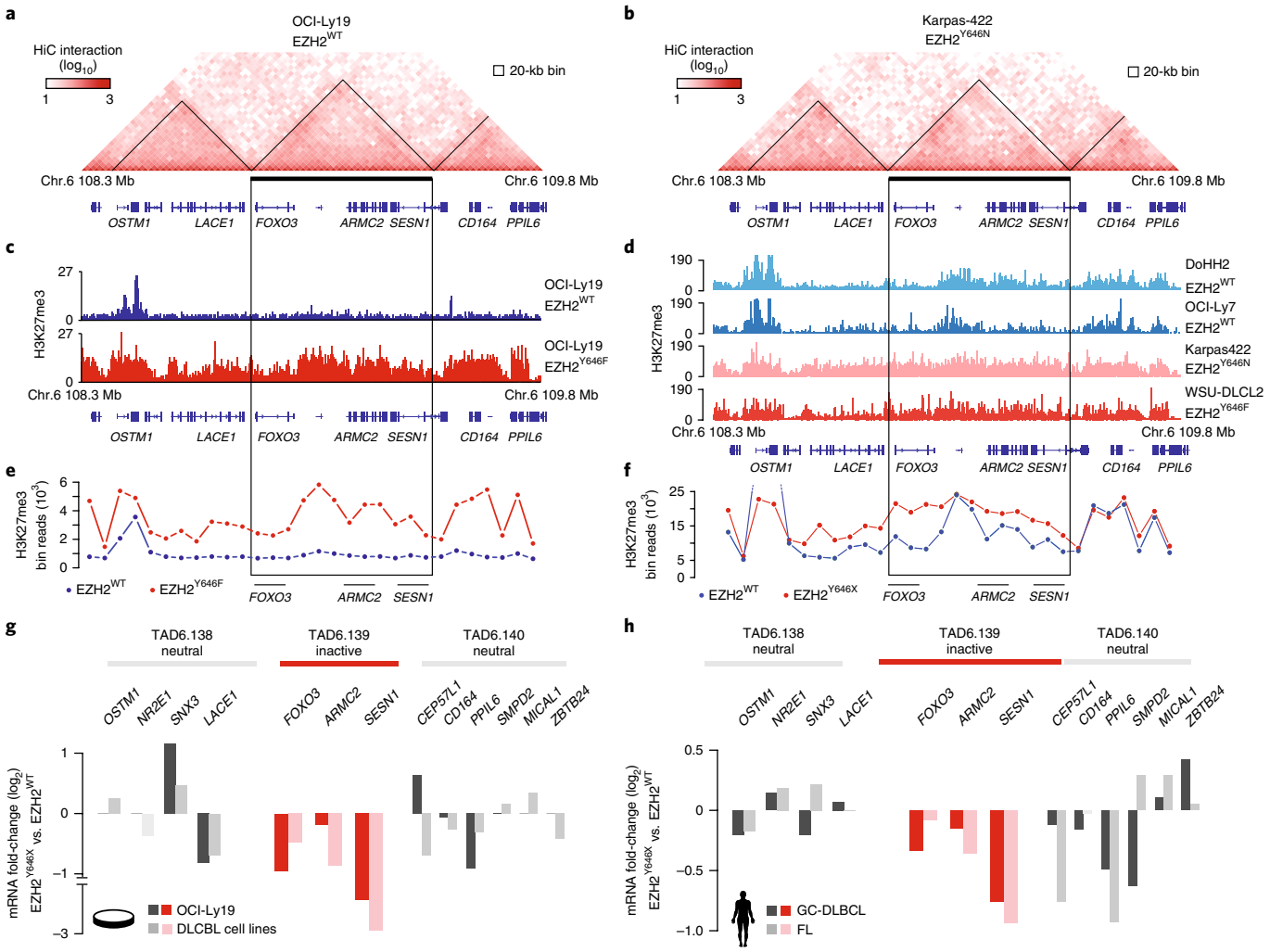


Fig. 3 | Inactivation of the tumor-suppressive TAD comprising *FOXO3* and *SESN1*. **a**,**b**, Hi-C contact maps at 20-kb resolution of the genomic region Chr.6 108.3–109.8 Mb in OCI-Ly19-EZH2^{WT} (**a**) and Karpas422-EZH2^{Y646N} (**b**). TADs are contoured in black, and representative genes within the TADs are shown. **c**,**d**, H3K27me3 ChIP-seq tracks on Chr.6 108.3–109.8 Mb for OCI-Ly19-EZH2^{WT} and OCI-Ly19-EZH2^{Y646F} ($n=3$ experiments per cell line) (**c**) and for the indicated EZH2^{WT} (DoHH2 and OCI-Ly7) and EZH2^{Y646X} (Karpas422 and WSU-DLCL2) cell lines ($n=2$ experiments per cell lines, except WSU-DLCL2 for which $n=1$) (**d**). **e**,**f**, Superimposition of H3K27me3 number of reads within 50-kb bins in Chr.6 108.3–109.8 Mb for OCI-Ly19-EZH2^{WT} and OCI-Ly19 EZH2^{Y646F} (**e**) and EZH2^{WT} and EZH2^{Y646X} (**f**) cells. **g**, mRNA expression fold changes of genes in Chr.6 108.3–109.8 Mb obtained by comparing OCI-Ly19-EZH2^{Y646F} and OCI-Ly19-EZH2^{WT} cell lines and EZH2^{Y646X} ($n=12$ samples) and EZH2^{WT} ($n=4$ samples) cell lines. **h**, mRNA expression fold changes of the genes in Chr.6 108.3–109.8 Mb obtained by comparing EZH2^{Y646X} ($n=7$) versus EZH2^{WT} ($n=30$) GC-DLBCL patient samples and EZH2^{Y646X} ($n=6$) versus EZH2^{WT} ($n=17$) FL patient samples. Fold changes corresponding to the inactive TAD are highlighted in red.

OCI-Ly19-EZH2^{WT} cells. TAD expression fold changes were moderately yet significantly anticorrelated with TAD H3K27me3 fold changes (Pearson's correlation = −0.31, $P=3.2\times10^{-21}$, Supplementary Fig. 3f). Moreover, TADs exhibiting an H3K27me3 fold change greater than two and an mRNA expression fold change less than negative two represented the majority (77%) of TADs exhibiting at least a twofold difference in both molecular features (Supplementary Fig. 3f), a result consistent with *EZH2* mutations increasing H3K27me3 and repressing transcription. Within the cell-line cohort including Karpas422 and WSU-DLCL2 (GSE40792), we confirmed a significant anticorrelation between TAD mRNA and H3K27me3 fold changes (Pearson's correlation = −0.23, $P=6\times10^{-25}$, Fig. 2k), and TADs exhibiting H3K27me3 fold changes greater than two and mRNA expression fold changes less than negative two represented 86% of the TADs exhibiting at least a twofold difference in both molecular features. This anticorrelation was even more pronounced when only TADs comprising the most differentially expressed genes were considered (Supplementary Fig. 3g,h). Overall, EZH2^{Y646X}-driven epigenetic and transcriptional changes within TADs were found to be significantly anticorrelated.

The interdependency between gene regulation and TAD structure prompted us to investigate TADs that exhibited a strong increase in H3K27me3 and a strong decrease in gene expression. Within the full set of conserved TADs, we identified 72 'inactive' TADs characterized by a greater than twofold increase in H3K27me3 and a greater-than-twofold decrease in mRNA expression (Fig. 2l). Notably, largely overlapping results were obtained when we considered OCI-Ly19 expression differences (EZH2^{Y646F} versus EZH2^{WT}) or Karpas422 and WSU-DLCL2 H3K27me3 profiles (Supplementary Fig. 3i). Inactive TADs were enriched in genes that were significantly downregulated in EZH2^{Y646X} cell lines (Fisher's exact test $P=2.8\times10^{-37}$, odds ratio (OR) = 12.4). Moreover, in 32 of these 72 domains, more than 80% of the genes within the domain exhibited concordant negative fold changes.

We defined these TADs as 'stringent inactive' (Fig. 2l) and the remaining 40 TADs as 'permissive inactive' (Fig. 2l). In primary patient samples, the mRNA fold changes in inactive TADs were significantly lower than those in other TADs (GC–DLBCL, two-tailed *t* test $P=2.3\times10^{-5}$ and FL, two-tailed *t* test $P=1.7\times10^{-8}$), and inactive TADs were enriched among TADs with concordant negative fold changes in more than 80% of the genes within their boundaries (Fisher's exact test: $P_{GC-DLBCL}=0.0006$, $OR_{GC-DLBCL}=3.2$, and $P_{FL}\,2.9\times10^{-5}$, $OR_{FL}=3.6$).

To explore the functional relevance of inactive TADs, we performed a gene-set enrichment analysis on all genes within these domains. Multiple Gene Ontology categories scored as significant (Supplementary Table 3), including cancer and lymphoma-specific categories (Fig. 2m). Notably, most significantly enriched gene sets contained EZH2 targets previously identified in prostate cancer[32] and genes differentially expressed in B-lymphocyte progenitors (Fig. 2m and Supplementary Table 3). These findings are consistent with $EZH2^{Y646X}$ locking in an inactive state genes that are transiently repressed in the GC during B-cell differentiation[21]. To determine whether inactive TADs are also repressed in normal GC cells, we compared the mRNA expression of their genes in normal centroblasts and centrocytes versus differentiated memory B cells[33]. Across all 72 inactive TADs, we identified seven domains that exhibited significant downregulation in centrocytes and centroblasts compared with memory B cells (Supplementary Fig. 3j). Transcriptionally repressed domains included genes regulating B-cell proliferation (for example *SESN1*)[22], DNA repair (*DTX3L* and *PARP9*)[34], and lymphocyte migration and trafficking (for example *S1PR1*)[35] (Supplementary Fig. 3k).

Overall, these results suggest that genes within inactive TADs are bona fide targets of EZH2 and that inactivation of TADs containing genes involved in B-cell differentiation and proliferation is potentially selected in EZH2-mutated tumors.

Silencing multiple tumor suppressors in inactive TADs. One of the top stringent inactive TADs included the genes *SESN1*, *FOXO3*, and *ARMC2* (TAD6.139 at chromosome (Chr.) 6 108850–109450 kb; Fig. 3a,b). Recently, we reported that *SESN1* is a direct target of mutated EZH2 and that it acts as a tumor suppressor in FL[22]. *FOXO3* has been described as bona fide tumor-suppressor gene controlling multiple signalings[36]. Moreover, genes within this locus are frequently co-deleted in non-Hodgkin lymphoma, and these deletions are largely mutually exclusive with *EZH2* mutations[22,37,38]. The inactive TAD6.139 exhibited a significant increase in H3K27me3 encompassing the entire domain in OCI-Ly19 cells overexpressing $EZH2^{Y646F}$, whereas H3K27me3 was detected only in specific regions in OCI-Ly19-$EZH2^{WT}$ cells (Fig. 3c,e). Similarly, an independent comparison of H3K27me3 levels in EZH2-mutated and wild-type cells confirmed a clear increase in H3K27me3 within the entire inactive TAD (Fig. 3d,f). Consistently, all genes within TAD6.139 exhibited lower mRNA expression in OCI-Ly19-$EZH2^{Y646F}$ and in EZH2-mutated lymphoma cells (Fig. 3g), as well as in GC-DLBCL and FL patient samples with $EZH2^{Y646X}$ mutations (Fig. 3h) compared with $EZH2^{WT}$ lymphomas. Enrichment in H3K27me3 was confirmed by targeted ChIP (Supplementary Fig. 4a), and changes in expression in OCI-Ly19-$EZH2^{Y646F}$ cells were confirmed by quantitative PCR (Supplementary Fig. 4b). Notably, additional genes with previously described tumor-suppressive function were found within inactive TADs (Supplementary Table 2). For example, similar patterns of H3K27me3 and gene downregulation were confirmed in TAD3.47 (Chr.3 33150–33450 kb), encompassing the tumor-suppressor gene *FBXL2* (refs. [39,40]) and silenced targets in lymphoma, *CRTAP*[41], and *SUSD5*[42] (Supplementary Fig. 4c–e). In particular, FBXL2 mediates degradation of cyclin-D3, which is the most expressed and mutated cyclin-D protein in B-cell lymphoma[43].

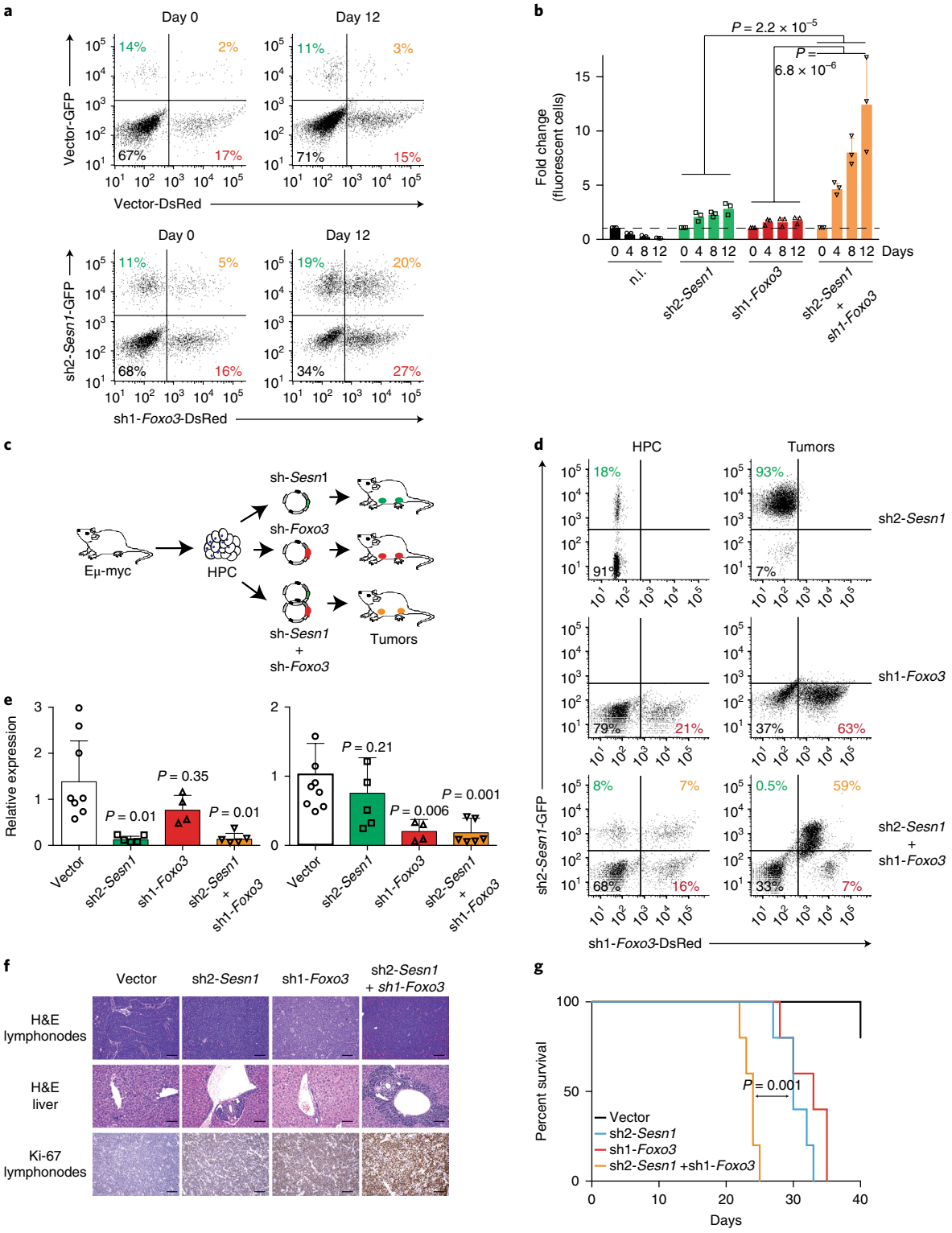
Together, these results indicate that mutated *EZH2* promotes concordant downregulation of multiple tumor suppressors within the same TAD, thus suggesting that these domains may function as tumor-suppressive units.

TAD inactivation accelerates lymphomagenesis. To assess the functional consequences of simultaneous downregulation of multiple tumor-suppressor genes within the same chromatin domain, we knocked down, alone and in combination, the expression of *FOXO3*, *ARMC2*, and *SESN1*, which are included in the candidate tumor-suppressive TAD6.139. We used IL3-dependent pro-B FL5-12 cells as a surrogate model to measure the effects of gene downregulation on B-cell proliferation[38,44]. We knocked down *Foxo3*, *Sesn1*, and *Armc2* by using short hairpin RNAs (shRNAs, denoted by 'sh'prefix) coupled with a fluorescent marker[22,44] (Supplementary Fig. 5a–c). The number of cells expressing either sh*Foxo3* or sh*Sesn1* increased two to threefold after four cycles of IL3 withdrawal (12 days). Notably, in the same time frame, cells with the dual knockdown of *Foxo3* and *Sesn1* increased 7- to 12-fold (Fig. 4a,b, Supplementary Fig. 5d, and Supplementary Table 4). We observed a similar enrichment in double-positive cells in normal growth conditions (Supplementary Table 4), whereas a synergistic effect was not observed with concurrent knockdown of *Armc2* and either *Foxo3* or *Sesn1* (Supplementary Fig. 5e,f), thus suggesting a specific functional synergy between loss of *Foxo3* and *Sesn1* to enhance pro-B cell proliferation.

Next, to test the oncogenic potential of the dual loss of *Foxo3* and *Sesn1* in tumor development and progression, we transduced hematopoietic progenitor cells (HPCs) isolated from an Eμ-myc mouse model[45,46] with shRNAs targeting *Foxo3*, *Sesn1*, or both (Fig. 4c). Animals harboring tumors expressing sh*Foxo3* ($n=16$), sh*Sesn1* ($n=17$), or both ($n=21$) had shorter overall survival than those expressing vector ($n=14$, sh*Foxo3* versus vector $P=0.04$, sh*Sesn1* versus vector $P=0.03$, and sh*Foxo3* and sh*Sesn1* versus vector $P=0.02$) (Supplementary Fig. 5g). Tumors that originated from HPCs transduced with single shRNAs were enriched in either

Fig. 4 | Concurrent downregulation of genes within tumor-suppressive TADs accelerates B-cell proliferation and lymphoma progression. a, Representative flow cytometry analysis of FL5-12 cells expressing empty vectors coupled with dsRed or GFP, or the sh1-*Foxo3*-dsRed and the sh2-*Sesn1*-GFP at days 0 and 12. Numbers indicate the percentages of cells in each subpopulation ($n=3$ independent experiments). **b**, Percentages of cells in each subpopulation at days 4, 8, and 12, relative to day 0. Bars indicate mean values, and error bars correspond to one standard deviation over three independent experiments. *P* values were calculated by two-way analysis of variance (independent variables are shRNA and day of measurement). **c**, Study design of in vivo experiments: HPCs were isolated from Eμ-myc embryos, retrovirally modified with the indicated plasmids, and transplanted into recipient animals. **d**, Representative flow cytometry analysis of HPC before transplantation (left column) and of isolated tumor cells (right column). Numbers indicate the percentages of cells in each subpopulation (sh2-*Sesn1* $n=3$, sh1-*Foxo3* $n=4$, sh2-*Sesn1*+sh1-*Foxo3* $n=4$, independent animals). **e**, Quantitative expression analysis of tumors expressing vector ($n=4$ independent animals and two technical replicates), sh2-*Sesn1* ($n=5$ independent animals), sh1-*Foxo3* ($n=4$ independent animals), and sh2-*Sesn1* and sh1-*Foxo3* ($n=6$ independent animals). Bars indicate mean values, and error bars correspond to one standard deviation. *P* values were calculated with two-tailed *t* test. **f**, Hematoxylin and eosin (H&E) cellular staining and Ki-67 immunohistochemistry of tumor and liver tissue biopsies of Eμ-myc mice expressing the indicated shRNAs or vector ($n=2$ independent experiments). Scale bar, 200 μm. **g**, Overall survival of secondary recipient animals transplanted with Eμ-myc tumor cells expressing vector ($n=5$), sh1-*Foxo3* ($n=5$), sh2-*Sesn1* ($n=5$) or sh2-*Sesn1* and sh1-*Foxo3* ($n=5$). *P* value was calculated with log-rank test.

GFP-positive (sh*Sesn1*) or dsRed-positive (sh*Foxo3*) cells (Fig. 4d). Notably, even though HPCs transduced with both shRNAs showed a low percentage of double-positive cells in the initial population (5–8%), tumors originating from these HPCs were almost exclusively composed of double-positive cells (Fig. 4d), thus suggesting that cells expressing both shRNAs expanded more rapidly than cells with a single shRNA. We confirmed that tumors expressing single or double shRNAs efficiently downregulated the expression of targeted genes (Fig. 4e), and all of the animals exhibited characteristic features of aggressive lymphoma (Fig. 4f and Supplementary Fig. 5h,i). To determine whether dual loss of *Foxo3* and *Sesn1* boosts tumor aggressiveness, we transplanted purified tumor cells with single and double knockdown into secondary recipient animals. Here, we observed a significant acceleration of tumor development ($P = 0.001$) (Fig. 4g) in animals receiving cells with double knockdown compared with those with a single shRNA. These results



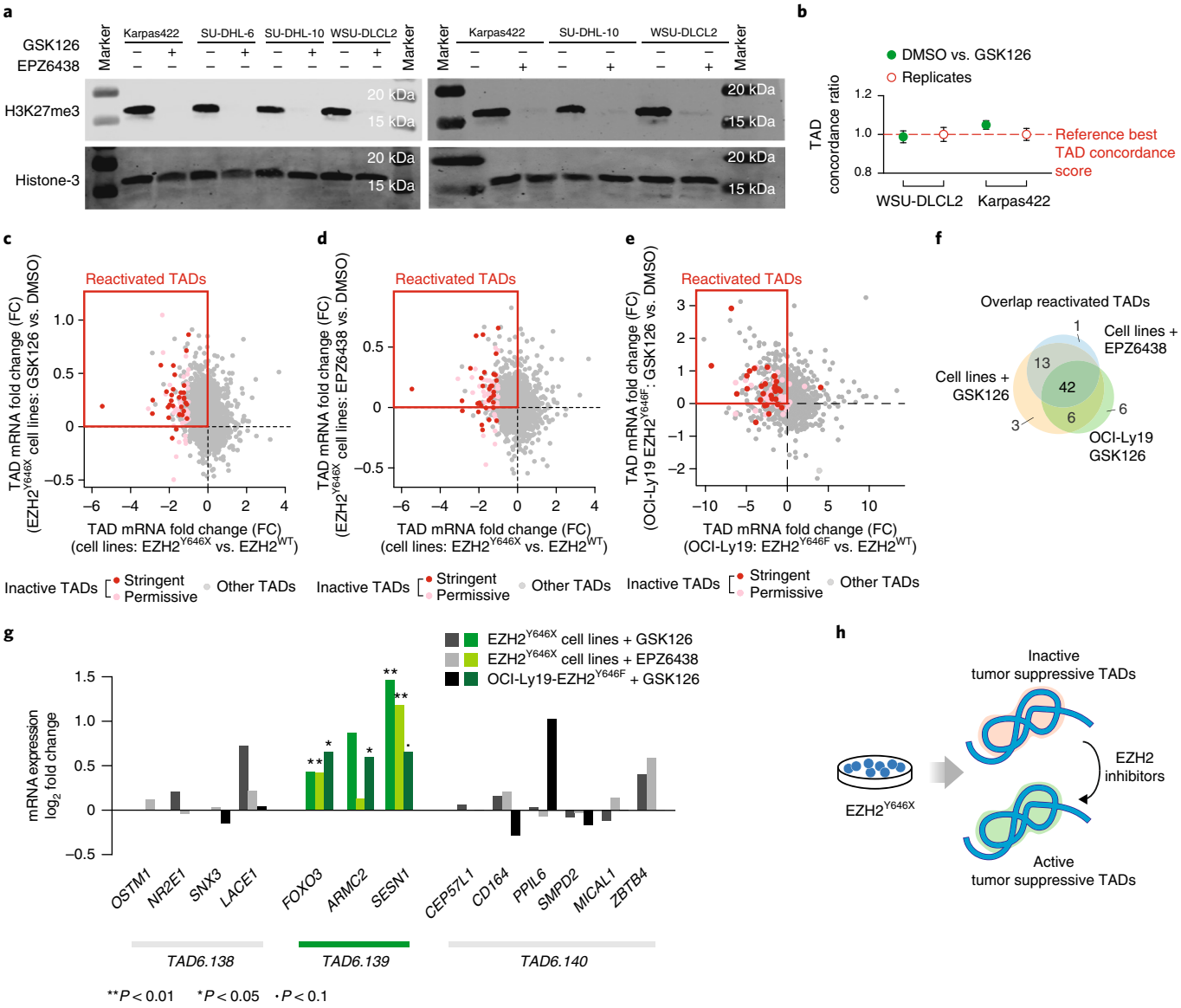


Fig. 5 | Inhibition of EZH2 abrogates H3K27me3 and reactivates inactive TADs. a, Immunoblot for H3K27me3 in EZH2^{Y646X} cell lines treated with 2 µM of GSK126 or 1 µM of EPZ6438 or vehicle (DMSO) for 72 h. Total histone H3 was used as a loading control, and the experiment was repeated twice. **b**, Ratio of TAD concordance (MoC) between WSU-DLCL2 and Karpas422 treated with DMSO or GSK126 and TAD concordance between independent replicates of each cell line (DMSO). Dots indicate the means of 22 intrachromosomal-map comparisons, and error bars are ±1 standard deviation. **c**,**d**, TAD mRNA expression fold changes in EZH2^{Y646X} ($n = 12$) versus EZH2^{WT} ($n = 4$) cell lines (*x* axis) versus TAD mRNA expression fold changes in EZH2^{Y646X} cells treated with GSK126 ($n = 12$) versus DMSO ($n = 12$) (**c**) or treated with EPZ6438 ($n = 12$) versus DMSO (n = 12) (*y* axis) (**d**). **e**, TAD mRNA expression fold changes of OCI-Ly19-EZH2^{Y646F} versus OCI-Ly19-EZH2^{WT} (*x* axis) and in OCI-Ly19-EZH2^{Y646F} treated with GSK126 versus DMSO (*y* axis). **f**, Intersection of inactive TADs obtaining a $\log_2(FC_{mRNA}) > 0$ (reactivated) in EZH2^{Y646X} cell lines treated with GSK126 or EPZ6438, or OCI-Ly19-EZH2^{Y646F} cells treated with GSK126. **g**, mRNA expression fold changes of individual genes in Chr.6 108.3–109.8 Mb obtained by comparing EZH2^{Y646X} cell lines treated with GSK126 ($n = 12$) or EPZ6438 ($n = 12$) versus DMSO ($n = 12$) and OCI-Ly19-EZH2^{Y646F} treated with GSK126 ($n = 3$) versus vehicle ($n = 3$). Fold changes corresponding to the inactive TAD are highlighted in green; *P* values were computed with limma two-sided *t* test. **h**, Graphical summary of the results: treatment with EZH2 inhibitors act as a switch turning tumor-suppressive TADs from transcriptionally inactive to active domains. *$P < 0.05$, **$P < 0.01$, •$P < 0.1$.

demonstrate that the inactivation of the tumor-suppressive TAD6.139 by mutated *EZH2* synergistically downregulates multiple tumor suppressors to drive tumor development and progression.

EZH2^{Y646X} inhibition reactivates transcription within TAD. Next, we explored the potential of blocking EZH2 methyltransferase activity by using two pharmacological inhibitors (GSK126 and EPZ6438) to reactivate inactive TADs. Pharmacological inhibition of EZH2 with either GSK126 or EPZ6438 efficiently depleted H3K27me3 in lymphoma cells (Fig. 5a). However, loss of H3K27me3 did not significantly modify the chromatin compartmentalization into TADs. Indeed, TADs derived from intrachromosomal maps of treated and untreated cells exhibited conservation scores similar to those observed for independent replicates of the same cell line (Fig. 5b). Nevertheless, EZH2^{Y646X} cell lines treated with the EZH2 inhibitors GSK126 and EPZ6438 ($n = 12$ treated and $n = 12$ DMSO) exhibited multiple transcriptional changes, although not as extreme as those observed between EZH2^{Y646X} and EZH2^{WT}

cells (Supplementary Fig. 6a,b). By mapping differentially expressed genes to our list of conserved TADs, we found that the TAD mRNA expression changes induced by the two inhibitors were positively correlated (Supplementary Fig. 6c) and mostly corresponded to upregulation (positive fold changes) of TADs that were downregulated (negative fold changes) in EZH2^{Y646X} compared with EZH2^{WT} cells (Supplementary Fig. 6d,e). Importantly, most inactive TADs (64/72 with GSK126 and 56/72 with EPZ6438) exhibited positive fold changes, and these reactivated TADs were enriched in significantly upregulated genes (adjusted P value < 0.01) from both experiments (Fig. 5c,d) (DMSO versus GSK126: Fisher's exact test $P = 10^{-4}$, OR = 2.74, DMSO versus EPZ6438: Fisher's exact test $P = 6.8 \times 10^{-8}$, OR = 3.17). A similar reactivation (54/72 inactive TADs) was verified in OCI-Ly19-EZH2^{Y646F} cells treated with the GSK126 inhibitor (Fig. 5e and Supplementary Fig. 6f). Reactivated TADs had a high degree of overlap among the three experiments (Fig. 5f). In particular, after treatment with both EZH2 inhibitors, genes within the tumor-suppressive TAD6.139 exhibited positive fold changes in all three models, results consistent with a stringent inactive-to-active switch of the domain, in which all genes are concordantly regulated (Fig. 5g). Conversely, TADs flanking TAD6.139 showed no pattern of co-regulation. These results were validated through quantitative PCR (Supplementary Fig. 6g). Notably, treatment with EZH2 inhibitors increased *FOXO3*, *SESN1*, and *ARMC2* expression exclusively in OCI-Ly19-EZH2^{Y646F} cells, whereas no significant expression changes were observed in the OCI-Ly19-EZH2^{WT} cells (Supplementary Fig. 6h). Reactivation of the tumor-suppressive TAD6.139 is thus a direct effect of inhibiting the mutated form of EZH2. Together, these results indicate that pharmacological depletion of H3K27me3 is sufficient to restore the transcriptional activity in previously silenced tumor-suppressive chromatin domains (Fig. 5h).

EZH2^{Y646X} modulates intra-TAD chromatin interactions. Changes in transcriptional activity are frequently associated with novel or missing interactions among regulatory elements[10,12,47]. Although TADs in EZH2^{Y646X}, untreated or treated with EZH2 inhibitors, and EZH2^{WT} cells were highly conserved (Supplementary Fig. 7a), we asked whether transcriptional changes in tumor-suppressive TADs were associated with rewiring of specific interactions within the domains.

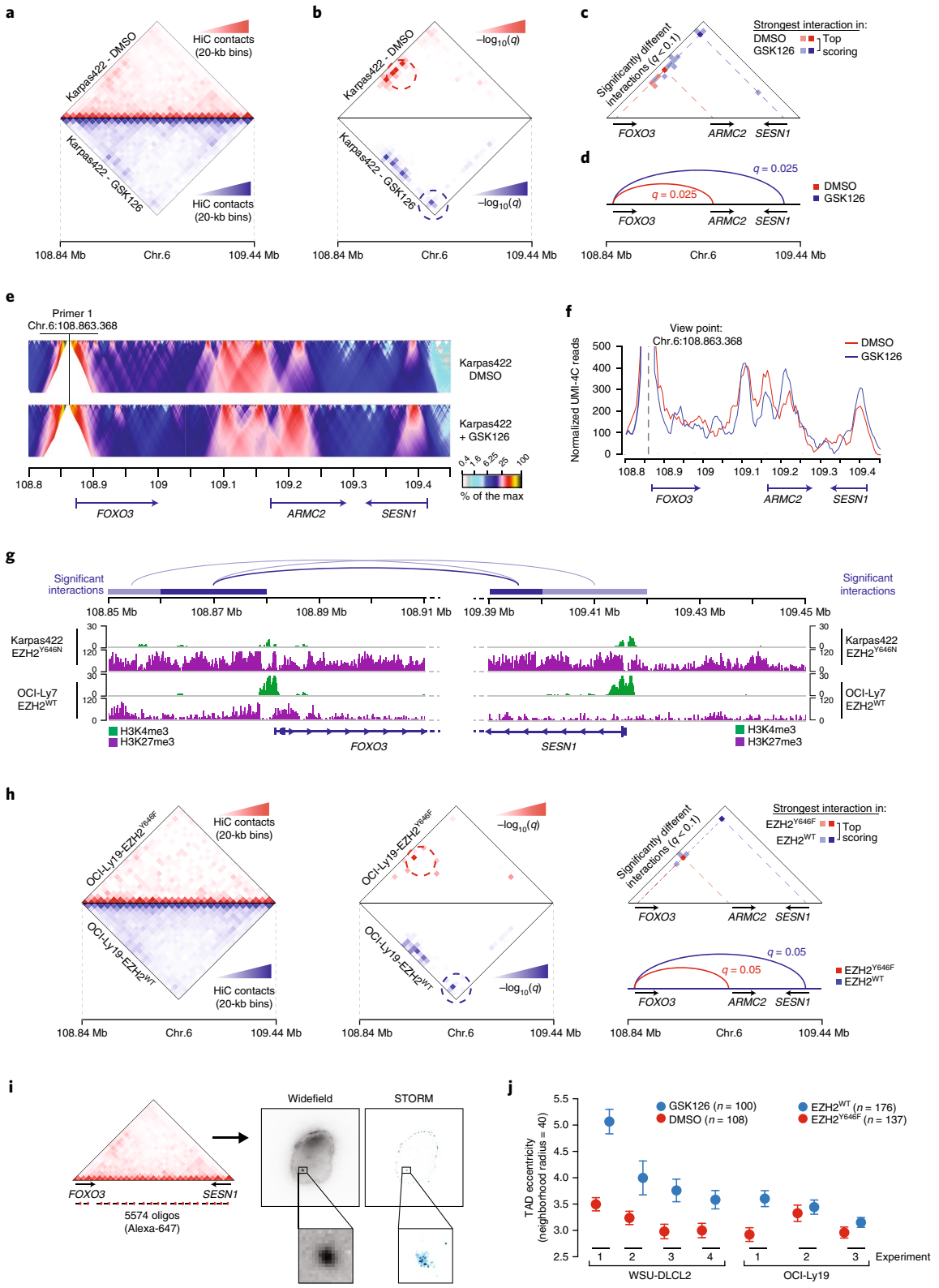
To address this question within the TAD6.139, we compared 20-kb-resolution Hi-C maps for Karpas422 treated with vehicle (Karpas422-DMSO, Fig. 6a) and with GSK126 (Karpas422-GSK126, Fig. 6a). Significant interactions between the 5' ends of the domain (Chr.6 108860–108880 kb) and a region spanning ~100 kb in the middle part of the TAD (Chr.6 109100–109220 kb) were detected in treated and untreated cells (Supplementary Table 5). However, only Karpas422 treated with GSK126 exhibited a highly significant interaction between genomic regions near the 5' and 3' ends of the domain (Chr.6 108860–108880 kb and 109380–109400 kb) (Fig. 6b). Next, we systematically compared significant interactions in Karpas422-DMSO and/or Karpas422-GSK126 to identify significant differences between these conditions (empirical q value < 0.1) (Fig. 6c). Notably, top-scoring differential interactions (empirical q value = 0.025) highlighted a stronger interaction between the 5' end and the middle region of the domain in Karpas422-DMSO than Karpas422-GSK126, whereas Karpas422-GSK126 exhibited a significantly stronger interaction between the 5' and 3' ends of the domain than Karpas422-DMSO (Fig. 6c,d and Supplementary Table 5). This observation was confirmed by quantification of the number of normalized reads between the 20-kb regions Chr.6 108860–108880 kb and Chr.6 109380–109400 kb ($P = 0.02$) (Supplementary Fig. 7b), and the same significantly different interactions were found in WSU-DLCL2 cells treated with vehicle or GSK126 (Supplementary Fig. 7c–e). To corroborate the observed changes in interaction frequency at higher resolution, we performed UMI-4C in Karpas422, either untreated or treated with GSK126, by using two independent sets of primers (Supplementary Table 6). We observed an enrichment in interactions between the Chr.6 108861–108863 kb and Chr.6 109370–109400 kb regions in cells treated with GSK126 compared with untreated cells (DMSO) with both primers (Fig. 6e,f and Supplementary Fig. 7f,g).

By mapping significantly differential interactions to functionally annotated regions in the genome, we found that these interactions linked genomic regions proximal to gene promoters that are recognizable by H3K4me3 peaks (Fig. 6g and Supplementary Fig. 7h). Notably, chromatin contacts were stronger between genomic regions proximal to *FOXO3* and *SESN1* promoters in EZH2^{Y646X} cells treated with GSK126 than in untreated cells (Fig. 6g). Concordant with the inactive status of TAD6.139 in EZH2^{Y646X} cells, the promoters of *FOXO3*, *ARMC2*, and *SESN1* exhibited lower H3K4me3 and higher H3K27me3 in EZH2^{Y646X} than in EZH2^{WT} cells (Fig. 6g and Supplementary Fig. 7h). Moreover, a peak of H3K27ac and H3K4me1 was detected at the *SESN1* promoter only in OCI-Ly7 EZH2^{WT} cells, thus suggesting that in these cells, this promoter might have distal enhancer function[48] (Supplementary Fig. 7h).

To determine whether treatment with EZH2 inhibitor restored intra-TAD interactions that were modified by the increase of H3K27me3 in EZH2 mutated tumors, we performed the same differential-interaction comparison between syngeneic EZH2^{Y646F} and EZH2^{WT} OCI-Ly19 cells as well as between Karpas422 (EZH2^{Y646N}) and OCI-Ly19 (EZH2^{WT}). OCI-Ly19-EZH2^{WT} showed stronger interactions between *FOXO3* and *SESN1* promoter regions than OCI-Ly19 cells expressing EZH2^{Y646F} (Fig. 6h). The same difference in interaction between these loci was detected between OCI-Ly19-EZH2^{WT} and Karpas422-EZH2^{Y646N} (Supplementary Fig. 7i). Conversely, both OCI-Ly19-EZH2^{Y646F} and Karpas422-EZH2^{Y646N}

Fig. 6 | Intra-TAD structural changes in EZH2^{WT} and EZH2^{Y646X} cells. a, Hi-C contact maps of TAD6.139 at 20-kb resolution (Chr.6 108.84–109.44 Mb) in Karpas422 treated with DMSO (top) or 2 μM GSK126 for 72 h (bottom). **b**, Significant interactions in Karpas422-DMSO (top) or Karpas422-GSK126 (bottom) determined by HiC-DC. Bin pairs within 2-Mb windows were tested. **c**, Significantly different interactions (q value < 0.1) between Karpas422-DMSO and Karpas422-GSK126 ($n = 435$ tested interactions). Empirical q values were determined as described in Methods. **d**, Most significantly different interactions between Karpas422-DMSO and Karpas422-GSK126. **e**, UMI-4C domainogram: mean number of contacts (percentage of the maximum) in the Chr.6 108.80–109.45 kb region. **f**, Normalized number of UMI-4C reads in Karpas422-DMSO and Karpas422-GSK126. **g**, Significantly stronger interaction (compare to **c**) in Karpas422-GSK126 than Karpas422-DMSO connecting Chr.6 108.85–108.81 Mb and Chr.6 109.39–109.45 Mb. H3K4me3 and H3K27me3 ChIP-seq tracks in Karpas422 and OCI-Ly7 ($n = 3$ experiments) within the same genomic coordinates. **h**, Left, Hi-C contact maps of TAD6.139 at 20-kb resolution in OCI-Ly19-EZH2^{Y646F} (top) and OCI-Ly19-EZH2^{WT} (bottom). Center, significant interactions in OCI-Ly19-EZH2^{Y646F} (top) or OCI-Ly19-EZH2^{WT} (bottom) determined by HiC-DC. Top right, significantly different interactions (q value < 0.1) between OCI-Ly19-EZH2^{Y646F} and OCI-Ly19-EZH2^{WT} ($n = 435$ tested interactions). Empirical q values were determined as described in Methods. Bottom right, most significantly different interactions between OCI-Ly19-EZH2^{Y646F} and OCI-Ly19-EZH2^{WT}. **i**, FISH library design for TAD6.139 and representative images acquired in wide field (left) and by STORM (right). **j**, Eccentricity of TAD6.139 (neighborhood radius = 40) in WSU-DLCL2 treated with 2 μM GSK126 or DMSO for 72 h ($n = 4$ experiments), and OCI-Ly19-EZH2^{WT} and OCI-Ly19-EZH2^{Y646F} ($n = 3$ experiments). Dots are means of multiple measurements of independent cells (Supplementary Table 7), and error bars are standard errors.

exhibited stronger interactions between promoter regions of *FOXO3* and *ARMC2* than OCI-Ly19-EZH2^{WT} (Fig. 6h and Supplementary Fig. 7i). To verify whether changes in promoter proximity also occurred in other inactive TADs, we performed a differential interactome analysis within the complete set of 72 inactive TADs, as well as within 100 sets of 72 randomly sampled neutral TADs. By comparing Karpas422-EZH2^{Y646N} and OCI-Ly19-EZH2^{WT} cells, we found 34 interactions that were significantly different within inactive TADs, but only 19 on average in neutral TADs ($P=0.04$, Supplementary Fig. 7j). Furthermore, significant different interactions between EZH2^{WT} and EZH2^{Y646X} cells were highly overlapping with those observed between cells treated with the EZH2 inhibitor GSK126 and untreated cells (17 out of 34, Fisher's exact test $P=6\times10^{-20}$, OR = 130) (Supplementary Fig. 7k). In summary, we



confirmed in multiple conditions, that the increase in H3K27me3 in *EZH2*-mutated tumors modulates promoter–promoter interactions within inactive TADs, which are coupled with silencing of gene transcription.

To investigate the degree of observable physical changes of the intra-TAD structure, we used a super-resolution imaging technique (stochastic optical reconstruction microscopy, or STORM)[49,50] and analyzed WSU-DLCL2 (DMSO versus GSK126, $n=4$ experiments) and OCI-Ly19 cells ($EZH2^{Y646F}$ versus $EZH2^{WT}$, $n=3$ experiments). To image the TAD structure by STORM, we designed a fluorescence in situ hybridization (FISH) library painting the entire TAD6.139 (Fig. 6i, Supplementary Fig. 8a, and Supplementary Note). TAD6.139 did not change significantly in size (as estimated by radius of gyration, R_g) between the two conditions in all experiments (Supplementary Table 7). However, we observed consistent changes in shape, as evidenced by higher eccentricity values in $EZH2^{WT}$ and $EZH2^{Y646X}$-GSK126 cells than in $EZH2^{Y646X}$ cells (Fig. 6j and Supplementary Fig. 8b–e). This increase in eccentricity was not observed in a neutral TAD of similar size (TAD1.54, Chr.1: 28650–29100 kb), exhibiting neither transcriptional nor epigenetic changes (Supplementary Fig. 8f,g).

Overall, our results indicate that the increase in H3K27me3 associated with *EZH2* gain-of-function mutations can modify the architecture and transcriptional activity within TADs.

Discussion

The recent advent of chromatin-conformation-capture techniques provides new opportunities and perspectives to unravel the effects of chromatin remodeling in cancer. $EZH2^{Y646X}$ increases H3K27me3 across the genome, thus leading to transcriptional repression. Here, we found that these epigenetic and transcriptional changes were not randomly distributed but were enriched and concordant within specific TADs, termed inactive TADs. Although future work will be needed to explore whether this specificity is predetermined or emerges through selection of oncogenic traits, inactive TADs were enriched in both EZH2 targets and B-cell differentiation programs, thus suggesting that both scenarios are likely and play a role in implementing $EZH2^{Y646X}$ oncogenic potential.

H3K27me3 determines changes in nucleosome proximity and is a marker of closed chromatin. Using multiple models, we showed that H3K27me3 modulates the proximity of gene promoters in association with gene expression changes. In particular, concordant expression changes of *FOXO3* and *SESN1* suggest co-regulated transcription. Interestingly, both genes are direct transcriptional targets of p53 (refs. [51,52]), thus prompting the hypothesis that promoter proximity may favor coordinated transcription of genes controlled by the same transcription factor.

Coordinated epigenetic and transcriptional reprogramming within TADs led to the synergistic inactivation of multiple tumor suppressors, enhancing B-cell proliferation and tumor growth. On the basis of these results, we introduced the concept of tumor-suppressive TADs. Importantly, analysis of expression changes in individual genes, rather than TADs, would probably have missed this synergistic effect. Moreover, although relatively small mRNA expression changes for individual genes are typically dismissed as random variations, the unexpected concordance of these variations within TADs could identify functionally relevant gene modules.

Chromatin domains can therefore provide a new lens through which to study oncogenic alterations, especially in the context of cancer epigenetic reprogramming.

URLs. ENCODE, https://www.encodeproject.org/; GEO, https://www.ncbi.nlm.nih.gov/geo/; GitHub custom scripts, https://github.com/CSOgroup/Donaldson-et-al-scripts/; Zenodo, https://zenodo.org/; NCBI BioProject https://www.ncbi.nlm.nih.gov/bioproject/; Bioconductor, https://bioconductor.org/; R, https://www.r-project.org/; Cufflinks, http://cole-trapnell-lab.github.io/cufflinks/; STAR, https://github.com/alexdobin/STAR/; Bowtie, http://bowtie-bio.sourceforge.net/bowtie2/manual.shtml; MACS2, https://github.com/taoliu/MACS/; Picard, http://broadinstitute.github.io/picard/; LIMMA, http://bioconductor.org/packages/release/bioc/html/limma.html; Bedtools, https://bedtools.readthedocs.io/en/latest/; HiC-DC, https://bitbucket.org/leslielab/hic-dc/; TopDom, http://zhoulab.usc.edu/TopDom/; mSIGdb, http://software.broadinstitute.org/gsea/msigdb/index.jsp; hg19 mappability file, http://hgdownload.soe.ucsc.edu/goldenPath/hg19/encodeDCC/wgEncodeMapability/wgEncodeCrgMapabilityAlign36mer.bigWig.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, statements of data availability and associated accession codes are available at https://doi.org/10.1038/s41588-018-0338-y.

Received: 27 April 2018; Accepted: 17 December 2018;
Published online: 28 January 2019

References

1. Bonev, B. & Cavalli, G. Organization and function of the 3D genome. *Nat. Rev. Genet.* **17**, 661–678 (2016).
2. Dixon, J. R., Gorkin, D. U. & Ren, B. Chromatin domains: the unit of chromosome organization. *Mol. Cell* **62**, 668–680 (2016).
3. Lieberman-Aiden, E. et al. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science* **326**, 289–293 (2009).
4. Dixon, J. R. et al. Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* **485**, 376–380 (2012).
5. Sexton, T. et al. Three-dimensional folding and functional organization principles of the Drosophila genome. *Cell* **148**, 458–472 (2012).
6. Nora, E. P. et al. Spatial partitioning of the regulatory landscape of the X-inactivation centre. *Nature* **485**, 381–385 (2012).
7. Le Dily, F. et al. Distinct structural transitions of chromatin topological domains correlate with coordinated hormone-induced gene regulation. *Genes Dev.* **28**, 2151–2162 (2014).
8. de Laat, W. & Duboule, D. Topology of mammalian developmental enhancers and their regulatory landscapes. *Nature* **502**, 499–506 (2013).
9. Fraser, J. et al. Hierarchical folding and reorganization of chromosomes are linked to transcriptional changes in cellular differentiation. *Mol. Syst. Biol.* **11**, 852 (2015).
10. Hnisz, D. et al. Activation of proto-oncogenes by disruption of chromosome neighborhoods.*Science* **351**, 1454–1458 (2016).
11. Weischenfeldt, J. et al. Pan-cancer analysis of somatic copy-number alterations implicates IRS4 and IGF2 in enhancer hijacking. *Nat. Genet.* **49**, 65–74 (2017).
12. Flavahan, W. A. et al. Insulator dysfunction and oncogene activation in IDH mutant gliomas. *Nature* **529**, 110–114 (2016).
13. Taberlay, P. C. et al. Three-dimensional disorganisation of the cancer genome occurs coincident with long range genetic and epigenetic alterations. *Genome Res.* **26**, 719–731 (2016).
14. Plass, C. et al. Mutations in regulators of the epigenome and their connections to global chromatin patterns in cancer. *Nat. Rev. Genet.* **14**, 765–780 (2013).
15. Morin, R. D. et al. Somatic mutations altering EZH2 (Tyr641) in follicular and diffuse large B-cell lymphomas of germinal-center origin. *Nat. Genet.* **42**, 181–185 (2010).
16. Souroullas, G. P. et al. An oncogenic Ezh2 mutation induces tumors through global redistribution of histone 3 lysine 27 trimethylation. *Nat. Med.* **22**, 632–640 (2016).
17. Tirode, F. et al. Genomic landscape of Ewing sarcoma defines an aggressive subtype with co-association of STAG2 and TP53 mutations. *Cancer Discov.* **4**, 1342–1353 (2014).
18. Comet, I., Riising, E. M., Leblanc, B. & Helin, K. Maintaining cell identity: PRC2-mediated regulation of transcription and cancer. *Nat. Rev. Cancer* **16**, 803–810 (2016).
19. Sneeringer, C. J. et al. Coordinated activities of wild-type plus mutant EZH2 drive tumor-associated hypertrimethylation of lysine 27 on histone H3 (H3K27) in human B-cell lymphomas. *Proc. Natl Acad. Sci. USA* **107**, 20980–20985 (2010).
20. Yap, D. B. et al. Somatic mutations at EZH2 Y641 act dominantly through a mechanism of selectively altered PRC2 catalytic activity, to increase H3K27 trimethylation. *Blood* **117**, 2451–2459 (2011).

21. Béguelin, W. et al. EZH2 is required for germinal center formation and somatic EZH2 mutations promote lymphoid transformation. *Cancer Cell* **23**, 677–692 (2013).
22. Oricchio, E. et al. Genetic and epigenetic inactivation of SESTRIN1 controls mTORC1 and response toEZH2 inhibition in follicular lymphoma. *Sci. Transl. Med.* **9**, eaak9969 (2017).
23. Rao, S. S. P. et al. A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell* **159**, 1665–1680 (2014).
24. Yang, T. et al. HiCRep: assessing the reproducibility of Hi-C data using a stratum-adjusted correlation coefficient. *Genome Res.* **27**, 1939–1949 (2017).
25. Pfitzner, D., Leibbrandt, R. & Powers, D. Characterization and evaluation of similarity measures for pairs of clusterings. *Knowl. Inf. Syst.* **19**, 361 (2009).
26. Shin, H. et al. TopDom: an efficient and deterministic method for identifying topological domains in genomes. *Nucleic Acids Res.* **44**, e70 (2016).
27. Zufferey, M., Tavernari, D., Oricchio, E. & Ciriello, G. Comparison of computational methods for the identification of topologically associating domains. *Genome. Biol.* **19**, 217 (2018).
28. Carty, M. et al. An integrated model for detecting significant chromatin interactions from high-resolution Hi-C data. *Nat. Commun.* **8**, 15454 (2017).
29. Jin, F. et al. A high-resolution map of the three-dimensional chromatin interactome in human cells. *Nature* **503**, 290–294 (2013).
30. McCabe, M. T. et al. EZH2 inhibition as a therapeutic strategy for lymphoma with EZH2-activating mutations. *Nature* **492**, 108–112 (2012).
31. Ortega-Molina, A. et al. The histone lysine methyltransferase KMT2D sustains a gene expression program that represses B cell lymphoma development. *Nat. Med.* **21**, 1199–1208 (2015).
32. Nuytten, M. et al. The transcriptional repressor NIPP1 is an essential player in EZH2-mediated gene silencing. *Oncogene* **27**, 1449–1460 (2008).
33. Klein, U. et al. Transcriptional analysis of the B cell germinal center reaction. *Proc. Natl Acad. Sci.* **100**, 2639–2644 (2003).
34. Yang, C.-S. et al. Ubiquitin modification by the E3 ligase/ADP-ribosyltransferase Dtx3L/Parp9. *Mol. Cell* **66**, 503–516.e5 (2017).
35. Matloubian, M. et al. Lymphocyte egress from thymus and peripheral lymphoid organs is dependent on S1P receptor 1. *Nature* **427**, 355–360 (2004).
36. Dansen, T. B. & Burgering, B. M. T. Unravelling the tumor-suppressive functions of FOXO proteins. *Trends. Cell Biol.* **18**, 421–429 (2008).
37. Okosun, J. et al. Integrated genomic analysis identifies recurrent mutations and evolution patterns driving the initiation and progression of follicular lymphoma. *Nat. Genet.* **46**, 176–181 (2014).
38. Oricchio, E. et al. The Eph-receptor A7 is a soluble tumor suppressor for follicular lymphoma. *Cell* **147**, 554–564 (2011).
39. Chen, B. B., Glasser, J. R., Coon, T. A. & Mallampalli, R. K. F-box protein FBXL2 exerts human lung tumor suppressor-like activity by ubiquitin-mediated degradation of cyclin D3 resulting in cell cycle arrest. *Oncogene* **31**, 2566–2579 (2012).
40. Li, L., Pan, D., Chen, H., Zhang, L. & Xie, W. F-box protein FBXL2 inhibits gastric cancer proliferation by ubiquitin-mediated degradation of forkhead box M1. *FEBS Lett.* **590**, 445–452 (2016).
41. Hatzimichael, E. et al. The collagen prolyl hydroxylases are novel transcriptionally silenced genes in lymphoma. *Br. J. Cancer* **107**, 1423–1432 (2012).
42. Chambwe, N. et al. Variability in DNA methylation defines novel epigenetic subgroups of DLBCL associated with different clinical outcomes. *Blood* **123**, 1699–1708 (2014).
43. Oricchio, E. et al. Frequent disruption of the RB pathway in indolent follicular lymphoma suggests a new combination therapy. *J. Exp. Med.* **211**, 1379–1391 (2014).
44. Mavrakis, K. J. et al. Genome-wide RNA-mediated interference screen identifies miR-19 targets in Notch-induced T-cell acute lymphoblastic leukaemia. *Nat. Cell Biol.* **12**, 372–379 (2010).
45. Scuoppo, C. et al. A tumour suppressor network relying on the polyamine-hypusine axis. *Nature* **487**, 244–248 (2012).
46. Schatz, J. H. et al. Targeting cap-dependent translation blocks converging survival signals by AKT and PIM kinases in lymphoma. *J. Exp. Med.* **208**, 1799–1807 (2011).
47. Lupiáñez, D. G. et al. Disruptions of topological chromatin domains cause pathogenic rewiring of gene-enhancer interactions. *Cell* **161**, 1012–1025 (2015).
48. Dao, L. T. M. et al. Genome-wide characterization of mammalian promoters with distal enhancer functions. *Nat. Genet.* **49**, 1073–1081 (2017).
49. Boettiger, A. N. et al. Super-resolution imaging reveals distinct chromatin folding for different epigenetic states. *Nature* **529**, 418–422 (2016).
50. Rust, M. J., Bates, M. & Zhuang, X. Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM). *Nat. Methods* **3**, 793–796 (2006).
51. Budanov, A. V. & Karin, M. The p53-regulated Sestrin gene products inhibit mTOR signaling. *Cell* **134**, 451–460 (2008).
52. Renault, V. M. et al. The pro-longevity gene FoxO3 is a direct target of the p53 tumor suppressor. *Oncogene* **30**, 3207–3221 (2011).

Acknowledgements

We thank B. Ren and A. D. Schmitt for help with Hi-C library preparation; C. Bolt for help with the UMI-4C protocol; J. Lingner, D. Trono, and F. Radtke for critical reading of the manuscript; and D. Duboule and J. Huelsken for useful discussions. We thank the EPFL research animal, flow cytometry, histology, and sequencing facilities. This work is supported by the ISREC Foundation (E.O.), the Swiss National Science Foundation (E.O. and M.C.D.-C. SNF-31003A_159637) and Swiss Cancer League (E.O. KFS-3982-08-2016). G.C. is supported by the Giorgi-Cavaglieri Foundation. D.T. is supported by the Swiss National Science Foundation (SNSF, SNF-310030_169519), M.Z. is supported by the Swiss Cancer League (KFS-3983-08-2016), and S.S. is supported by European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement no. 665667.

Author contributions

M.C.D-C. prepared Hi-C, UMI-4C, ChIP-seq and RNA-seq libraries and performed in vitro validation experiments and DNA FISH experiments. S.S. analyzed Hi-C, UMI-4C, and ChIP-seq data. M.Z. performed the comparison of Hi-C contact maps, TAD calling, and all analyses based on mRNA expression data. D.T. performed interactome analyses and all analyses based on ChIP-seq data. K.M.D. and T.R. acquired and analyzed STORM images. N.K. and E.B. performed in vitro and in vivo experiments. M.M. assisted in the analysis of Hi-C data. F.R. assisted in the analysis of STORM data. S.M. supervised in STORM acquisition and image analyses. G.C and E.O. designed the study and wrote the manuscript with comments from all authors.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41588-018-0338-y.

Reprints and permissions information is available at www.nature.com/reprints.

Correspondence and requests for materials should be addressed to G.C. or E.O.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Methods

Data generation and models. Lymphoma cell lines used in this study (OCI-Ly19, DoHH2, SU-DHL-10, WSU-DLCL2, Toledo, SU-DHL-6, and Karpas422) were authenticated within the past year by short tandem repeat (STR) cell-line authentication analysis (MicroArray, Switzerland) and cultured as described in the Supplementary Note. OCI-Ly19 cells were transduced with lentiviral particles to express EZH2Y646F in tandem with GFP or the empty vector (OCI-Ly19-EZH2^{Y646F}). After the initial transduction, 3–5% of the GFP-positive cells were sorted to obtain a pure population of GFP-positive cells. We confirmed by RNA-sequencing (RNA-seq) analysis that only one copy of mutated *EZH2* was integrated in the genome.

Hi-C experiments were performed on Karpas422, WSU-DLCL2, OCI-Ly19-EZH2^{WT}, and OCI-Ly19-EZH2^{Y646F}, treated with 2 µM GSK-126 or DMSO for 72 h. Data generation and analysis are detailed in the Supplementary Note. Targeted chromosome conformation capture using unique molecular identifiers (UMI-4C) was performed in Karpas-422 treated with 2 µM GSK126 or DMSO for 72 h, as described in ref. [53]. Hi-C and UMI-4C library preparation, contact-map generation, and data normalization were performed as described in the Supplementary Note. A list of the primers used in this study is provided in Supplementary Table 8.

mRNA expression was assessed by means of RNA-seq in OCI-LY19-EZH2^{WT}, OCI-Ly19-EZH2^{Y646F}, and OCI-Ly19-EZH2^{Y646F} after treatment with the EZH2 inhibitor GSK126 for 72 h. ChIP-seq was performed in OCI-Ly19-EZH2^{WT} and OCI-Ly19-EZH2^{Y646F}, Toledo, and WSU-DLCL2 cells by using a monoclonal antibody to H3K27me3 (Cell Signaling, 9733). H3K27me3 ChIP-seq data for OCI-Ly7, DOHH2, and Karpas422 were obtained from ENCODE, whereas for WSU-DLCL2, data were obtained from the GEO Database (GSE40792). Targeted ChIP validation was performed for *FOXO3* and *SESN1* on Toledo and WSU-DLCL2 cell lines. All data generation and analyses are detailed in the Supplementary Note.

STORM was used to compare the structure of TAD6.139 between WSU-DLCL2 treated with either 2 µM GSK126 or DMSO (four independent experiments), and between OCI-Ly19-EZH2^{WT} and OCI-Ly19-EZH2^{Y646X} (three independent experiments). The structure of the neutral TAD1.54 was compared between WSU-DLCL2 treated with either 2 µM GSK126 or DMSO (two independent experiments). STORM data were generated and analyzed as described in the Supplementary Note.

Mouse model. We isolated Eµ-myc transgenic fetal liver cells from embryos at embryonic day 13.5. The HPCs were grown for 4 d in a specially adapted growth medium (DMEM, IMDM, L-glutamine, pen/strep, FBS, and SCS-stem cell supplement (Whei, IL3, IL6, SCF, and polybrene) and retrovirally transduced with either the vector MLS-GFP containing shRNA for *Sesn1* or *Foxo3* or the empty vectors. We inoculated genetically modified HPCs into lethally irradiated, syngeneic wild-type recipient animals. Disease onset was monitored by palpation and blood smears. Data were analyzed in Kaplan–Meier format by using the log-rank (Mantel–Cox) test for statistical significance. Transgenic animals were maintained in EPFL SPF animal facility, and all animal studies were approved by Swiss Cantonal authorities (animal license VD2932 and VD2932.1).

Comparison of intrachromosomal Hi-C contact maps. All pairwise comparisons between intrachromosomal contact maps (or matrices) were based on four different metrics: overall matrix correlation by SCC, fraction of corresponding bins assigned to the same compartment (A or B), conservation of TADs, and conservation of significant interactions between bins.

Matrix correlation (stratum-adjusted correlation coefficient). To compare the overall distribution of contacts across intrachromosomal Hi-C maps, we used the SCC, as implemented in the *hicrep* R package[24] in Bioconductor (*get.scc* function, after having trained the smoothing parameter with the *htrain* function for the range 0:20). After smoothing Hi-C matrices with a 2D mean filter to reduce noise and bias effects, we stratified chromatin contacts according to the genomic distance between interacting loci. The number of strata K is given by the maximal interaction distance divided by the bin size (for example $K = 3{,}840$, for a 192-Mb chromosome binned at 50 kb; here, we chose the chromosome size as the maximal interaction distance). The SCC was then computed as a weighted average of stratum-specific Pearson's correlation coefficients, and it can be interpreted as a classical correlation coefficient with values ranging from −1, for perfect anticorrelation, to +1, for perfect correlation.

Fraction of corresponding bins assigned to the same compartment. Compartments were calculated as previously described[3]: Hi-C contact matrices were binned at 1 Mb by using the *pca.hic* function from the *HiTC* R package in Bioconductor, and compartments A and B were defined according to the signs of the values of the first eigenvector retrieved from principal component analysis of the Pearson correlation matrix of the observed over expected Hi-C matrix (when the first eigenvector separated the two chromosome arms, the second eigenvector was used). To compare the compartments between two intrachromosomal matrices, we defined the sign (positive or negative) associated with a specific compartment (A or B) as the one achieving the highest correlation between the eigenvectors of the two datasets. For each pair of 1,000-kb bins corresponding to the same genomic region in the two Hi-C contact matrices, we compared the sign of the corresponding eigenvalue: concordant signs corresponded to bins assigned to the same compartment, opposite signs corresponded to bins in different compartments. A similarity score for each pair of intrachromosomal maps was defined as the fraction of bins that were assigned to the same compartment.

Conservation of TADs. First, TADs were called at 50-kb resolution by using the TopDom[26] (*TopDom* function from the *TopDom* R package with a window size $w = 5$).

Next, we compared the set of TADs identified in two intrachromosomal matrices by computing the concordance of the two corresponding chromosome partitions. For this purpose, we used the measure of concordance (MoC) previously introduced to compare clustering assignments[25]. MoC is defined as follows:

$$MoC(\boldsymbol{P},\boldsymbol{Q}) = \begin{cases} 1 \ if \ N_P = N_Q = 1 \\ \dfrac{1}{\left(\sqrt{N_P N_Q} - 1\right)} \left(\sum_{i=1}^{N_P} \sum_{j=1}^{N_Q} \dfrac{\|F_{i,j}\|^2}{\|P_i\| \, \|Q_j\|} \right) otherwise \end{cases}$$

where **P** and **Q** are the sets of TADs being compared, including N_P and N_Q numbers of TADs, respectively. P_i and Q_j are two individual TADs within **P** and **Q** having size $||P_i||$ and $||Q_j||$, respectively, measured in base pairs. Finally, $||F_{ij}||$ corresponds to the size (number of base pairs) of the overlap between the two TADs P_i and Q_j.

Conservation of significant interactions. The statistical significance of chromatin interactions at the bin level was estimated with HiC-DC[28] in 'fixed bin' mode. The bin size was set to 50 kb, with degrees of freedom (df) = 6 and size = 1.0. The hg19 mappability file was retrieved by following the link provided in the URLs section. Interactions were called significant if the HiC-DC q value was smaller than 0.05. Conservation of significant interactions between two given cell lines c1 and c2 was computed as the overlap between their set of significant interactions as follows:

$$overlap = \frac{\#\ of\ common\ s.i.}{\min\{(\#\ of\ s.i.\ in\ c1), (\#of\ s.i.\ in\ c2)\}}$$

Generation of Hi-C contact maps for replicates and randomized maps for comparison. To build a reference scale of values for the different metrics, we included comparisons between our Hi-C maps and randomized Hi-C maps, as well as comparisons between independent replicates of the same cell line.

In the comparisons based on the SCC measure and on the fraction of conserved significant interactions, we generated randomized Hi-C maps by shuffling count values within each diagonal of the Hi-C matrix, that is by permuting the number of contacts among bins located at the same genomic distance. For the comparison of TAD by MoC, random TAD partitions were generated by shuffling the original set of TADs in a way that preserved the number and the size of the TADs. The resulting random maps were used to estimate the expected similarity score between Hi-C experiments performed without any expected structural similarity.

Additionally, for each cell line, we selected and aggregated independent replicates in two groups with similar total number of reads and generated separate contact maps and the list of contacts (interactomes) for each group. The maps and the interactomes were used to estimate the expected similarity score between Hi-C experiments performed on the same cell line (that is, expected highest similarity).

Concordance analyses of H3K27me3 and transcriptional changes within TADs. A consensus list of TADs conserved among Karpas422, OCI-Ly19-EZH2^{WT}, OCI-Ly19-EZH2^{Y646F}, and WSU-DLCL2 cell lines was determined by using an approach based on weighted interval scheduling (WIS) efficiently solved with a dynamic programming (DP) approach[54]. Briefly, each interval/TAD received a weight according to how conserved it was across cell lines (allowing a tolerance radius of two bins or 100 kb), and the goal was to identify a set of nonoverlapping intervals of maximum total weight (details in Supplementary Note). Using this approach, we identified a total of $n = 3{,}773$ consensus TADs.

H3K27me3 comparison among subcompartments. Chromosome subcompartments derived in GM12878 were downloaded from the GEO database (GSE63525). Subcompartment calls were used to partition contact maps generated by our experiments, given the high concordance between A and B compartments derived in GM12878 and in our cell lines (Fig. 1c,d). Moreover, subcompartments have been called only in GM12878, given that the procedure requires an extremely high number of reads (documentation in GSE63525). To compare H3K27me3 across subcompartments, we considered the mean H3K27me3 levels within each continuous DNA sequence that was assigned to the same subcompartment (sc region) and compared the distribution of mean H3K27me3 values among sc regions annotated for the same compartment. Unannotated sc regions were excluded from this analysis. H3K27me3 fold changes were computed for each sc

region through comparing the mean H3K27me3 values in OCI-Ly19-EZH2^{Y646F} versus OCI-Ly19-EZH^{WT} cells.

Intra-TAD and inter-TAD correlation between H3K27me3 fold changes and levels. Correlations between normalized H3K27me3 profiles of OCI-Ly19-EZH2^{Y646F} and OCI-Ly19-EZH2^{WT} cells were investigated as previously described[23]. Briefly, each TAD was divided into ten equally spaced bins, as well as two flanking regions of the same size of the TAD. Mean H3K27me3 levels and the fold changes of mean H3K27me3 levels in mutated versus wild-type samples within each of these 30 bins were then stored as a row of an $N \times 30$ matrix, where N is the number of TADs. Finally, Spearman correlation coefficients were computed between each pair of columns of this matrix, thus constructing a 30×30 symmetric pairwise correlation matrix, whose columns 11–20 corresponded to bins within a TAD, whereas columns 1–10 and 21–30 corresponded to bins within upstream and downstream regions, respectively. The central 10×10 block represented intra-TAD correlations, whereas the 10×10 blocks above (or below) it and on its right (or left) corresponded to inter-TAD correlations. The correlation of H3K27me3 fold changes as a function of distance for intra-TAD and inter-TAD pairs of loci (Fig. 2d) was performed as follows:

1. Each TAD was divided into bins of 25 kb, as well as its two flanking regions (of the same size as the TAD) upstream and downstream; fold changes in each bin were computed as explained above.
2. A distance vector was constructed as a sequence of distances ranging from 25 kb to 1 Mb, each of them separated by 25 kb.
3. For each element i of the distance vector, we extracted all pairs of loci that were separated by a distance equal to i; next, we derived an $N \times 2$ matrix storing the N pairs for which both loci were in the same TAD, and we computed the Spearman correlation coefficient between the two columns, which represented the intra-TAD correlation at distance i; analogously, we derived an $M \times 2$ matrix storing the M pairs for which one locus was in the TAD, and the other was in one of the two flanking regions and computed the inter-TAD correlation for distance i in the same way.
4. Finally, we plotted the intra-TAD and inter-TAD correlation versus distance in two different lines.

Concordance between H3K27me3 changes and TADs. H3K27me3 ChIP-seq levels in OCI-Ly19-EZH2^{Y646F} cells and in OCI-Ly19-EZH2^{WT} were jointly normalized across the two cell lines, as described in the Supplementary Note. Normalized H3K27me3 levels were averaged within each conserved TAD, and the $\log_2$ fold change was computed between the two conditions (mutated versus wild type). Density distributions of $\log_2$ fold change values were visualized by using the 'density' R function.

Random distributions of H3K27me3 TAD fold changes were generated by permutation of intervals (bins) of H3K27me3 read counts. Specifically, for a given size S, each chromosome was partitioned into bins of S base pairs, and read counts within each bin were summed. Corresponding bins in EZH2^{Y646F} and EZH2 wild-type OCI-Ly19 cells were permuted together to preserve the H3K27me3 ratio between the two conditions for each bin. Bins within each chromosome were permuted separately. TAD fold changes were recalculated as described above. Random distributions were generated with bin sizes ranging from 1 kb to 2 Mb. The number of TADs obtaining a fold change greater than two was compared when using the observed H3K27me3 distribution and distributions generated with bin size ranging from 50 kb to 2 Mb.

Concordance between transcriptional changes and TADs. For each mRNA expression dataset, genes with detectable expression were assigned to the TADs with the highest overlap, and genes nonoverlapping with a TAD were discarded. TADs containing fewer than three genes or more than the 0.99-quantile of the number of genes by TAD distribution were excluded. With these filters, we retained $n = 2{,}038$ TADs for the comparison between EZH2^{Y646X} and EZH2^{WT} DLBCL primary samples and cell lines, $n = 1{,}923$ TADs for the comparison of follicular lymphoma primary samples, and $n = 900$ TADs for the comparison of OCI-Ly19 EZH2^{Y646F} and OCI-Ly19-EZH2^{WT}. For the latter comparison, we retained considerably fewer TADs, a result consistent with the RNA-seq dataset including only three replicates for OCI-Ly19-EZH2^{Y646F} and three replicates for OCI-Ly19-EZH2^{WT}; thus, fewer genes had detectable expression in both conditions than in the other datasets.

To test whether the mRNA expression changes were more concordant within TADs than expected, we first defined TAD mRNA expression fold changes as the mean of the fold changes of each gene within the TAD. Gene fold changes were determined for both microarray and RNA-seq datasets as described above. Next, we defined a measure of concordance for fold changes within a TAD and compared the observed cumulative sum distribution of concordance values obtained by our set of conserved TADs with the expected cumulative sum distribution under random permutation of the gene-to-TAD assignments. Briefly, an FCC score was computed for each TAD as:

$$FCC = \left(2\frac{\#FC^-}{\#FC} - 1\right) * \left(2\frac{\sum |FC^-|}{\sum |FC|} - 1\right)$$

where #FC is the number of genes within the TAD, #FC$^-$ is the number of genes with a negative fold change within the TAD, |FC| are absolute fold-change values, and |FC$^-$| are absolute values of negative fold changes ($\log_2$ fold changes were used). TADs were ranked from highest to lowest FCC, and the cumulative sum curve was calculated. For each test, we compared the observed cumulative sum curve against curves obtained after random permutation of gene-to-TAD assignments. Intuitively, the more concordant the gene expression changes within TADs, the steeper the increase in the cumulative sum. We considered the observed concordance to be greater than expected when the observed curve was higher than all random curves. To permute gene-to-TAD assignments, we divided genes in five classes according to their overall levels of expression and shuffled TAD labels within each class. For each test, we performed 10,000 permutations and reported the range comprising all random curves.

Pearson's correlation of mRNA expression between genes located in the same TAD or different TADs were compared as a function of their genomic distance (maximum 500 Kb). Microarray data were normalized by using the median absolute deviation method, whereas RNA-seq data was transformed by using quantile normalization. Curves were fitted by using the R *loess* function (from the *stats* package, with default parameters).

Genes within inactive TADs were analyzed by gene-set enrichment analysis, and their expression was assessed in centrocytes, centroblasts, and memory B-cells, as described in the Supplementary Note.

Differential intra-TAD interactome analysis for TAD6.139. For each pair of cell lines/conditions, significantly different interactions between 20-kb bins within the TAD 6.139 were determined as follows:

1. HiC-DC was run on Hi-C read pairs of Chr.6 with a bin size of 20 kb, and the q values obtained for each pair of bins within the entire chromosome were retained;
2. For each pair of bins i and j (with $i_s > j_s$), the interaction difference d between condition A and condition B was computed as $d_{i,j}^{A-B} = -\log_{10}(q_{i,j}^{A}) + \log_{10}(q_{i,j}^{B})$;
3. Distributions of $d_{i,j}^{A-B}$ values were determined for pairs of bins at the same genomic distance;
4. For each pair of bins i_s and j_s in TAD6.139, the significance of the corresponding interaction difference was tested by computing its empirical P value against the distribution of interaction differences computed between bins at the same distance, namely $\{d_{i,j}^{A-B}\}_{i,j|i-j=i_s-j_s}$;
5. Empirical P values obtained for all pairs of bins within TAD 6.139 were corrected for multiple testing by using the Benjamini–Hochberg procedure;
6. Finally, interactions were called significantly stronger in condition A than in condition B if they satisfied both the following conditions:
   a. The q value computed by HiC-DC was smaller than 0.1,
   b. The q value resulting from step (5) was smaller than 0.1.

Differential-interaction analysis for intra-TAD promoter-proximal regions. For each pair of cell lines/conditions A and B, significantly different interactions between 50-kb bins mapping to gene promoters located in a set of TADs were determined as follows:

1. HiC-DC was run on Hi-C read pairs of all chromosomes with a bin size of 50 kb, and the q values obtained for each pair of bins were retained;
2. Bin-level promoter-promoter interactions (PPIs) between genes x and y were identified as $PPI_{x,y} = -\log_{10}(q_{i,j})$, where i and j are 50-kb bins overlapping the transcription start sites of genes x and y, respectively. Only PPIs between genes in the same TAD and less than 2 Mb apart were retained;
3. Significantly different PPIs between conditions A and B were determined as described in steps 2–6 described in 'Differential intra-TAD interactome analysis', with a cutoff of 0.25 for q values computed in step 5

Random sets of neutral TADs were sampled from our consensus list of 2,038 TADs such that:

1. Only TADs with $|log_2(\text{H3K27me3 fold change}_{OCI-Ly19\ EZH2-Y646F\ vs\ EZH2-wt})| < 0.5$ and $|log_2(\text{mRNA fold change}_{Cell\ lines\ EZH2-Y646X\ vs\ EZH2-wt})| < 0.5$ were sampled;
2. The distributions of the numbers of genes per TAD in the random sets were the same as in the set of 72 inactive TADs

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

In this study, we used the following mRNA expression datasets: GSE23501 for wild-type and *EZH2*-mutated GCB-DLBCL primary human samples, PRJNA278311 (NCBI-BioProject) for wild-type and *EZH2*-mutated FL primary human samples, GSE40792 for wild-type and *EZH2*-mutated cell lines before and after treatment with GSK126, GSE49284 for *EZH2*-mutated cell lines before and after treatment

with EPZ6438, and GSE12195 for centrocytes, centroblasts, and memory B cells. ChIP-seq data for H3K27me3 in OCI-Ly7, DOHH-2, and Karpas-422 were downloaded from ENCODE; H3K27me3 in WSU-DLCL2 was downloaded from GSE40970; H3K4me3 in OCI-Ly7 and Karpas-422 were downloaded from ENCODE. ChIP-seq data for H3K27me3 and RNA-seq data for OCI-Ly19 and OCI-Ly19-EZH2^{Y646F} were generated as described in the manuscript and have been deposited at GSE114270. HiC matrices and UMI-4C data have been deposited at Zenodo: https://doi.org/10.5281/zenodo.1244182. Custom scripts are available through a public GitHub repository at: https://github.com/CSOgroup/Donaldson-et-al-scripts/.

References

53. Schwartzman, O. et al. UMI-4C for quantitative and targeted chromosomal contact profiling. *Nat. Methods* **13**, 685–691 (2016).
54. Kleinberg, J. & Tardos, É. *Algorithm Design* (Pearson, Boston, 2005).

nature research

Corresponding author(s): Elisa Oricchio, Giovanni Ciriello

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistical parameters

When statistical analyses are reported, confirm that the following items are present in the relevant location (e.g. figure legend, table legend, main text, or Methods section).

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	An indication of whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistics including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated
☐	☒	Clearly defined error bars *State explicitly what error bars represent (e.g. SD, SE, CI)*

Our web collection on statistics for biologists may be useful.

Software and code

Policy information about availability of computer code

Data collection	Data collected from public repository (e.g. ENCODE) was directly retrieved from the associated website without the use of custom software as described in the Methods section. Flow cytomentry data were acquired using LSR II (BD). qPCR data were acquired using AB Applied STEPOne PLUS. sequencing was performed using Illumina HI-seq and NexSeq500. Histology and immunofluorescence picture were acquired with Leica DM4000 B-LED with Leica application suite v2.0. Western blot were acquired with LICOR and Image Studio light program v3.1.4 STORM imaging were acquired using a custom build microscope (Douglass et al. Nat.Photonics 2016) but with a photometrics prime sCMOS camera.
Data analysis	Publicly available computational tools to derive TADs, compute the interactome, and process Hi-C data matrices were used as described in the Methods section and they include: Bowtie2-2.2.9 SAMtools (v1.3.1 and v1.6) Picard tools (v2.5.0) bwa mem (v0.7.12) STAR 2.5.3a

Cufflinks 2.2.1
SRAtools
R package Rsubread 1.28.1
R package limma 3.34.8
SICERpy (wrapper for SICER_V1.1)
IGV genome browser
bedGraphToBigWig v4
R package hicrep v1.0.0
TopDom v0.0.2
R package HiC-DC v1.0
mSigDB
PRISM 6
FlowJo 10.4.1
Custom scripts have been implemented in bash, R (v3.4.3) and Python (v2.7.5) and are made available through a public GitHub repository at: https://github.com/CSOgroup/Donaldson-et-al-scripts
Finally, standard statistical analyses and related figures were generated using standard R functions as described in the Methods section.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

In this study, we used the following mRNA expression datasets: GSE23501 for wild type and EZH2 mutated GC-DLBCL primary human samples and PRJNA278311 for follicular lymphoma primary human samples (Fig. 2), GSE40972 for wild type and EZH2 mutated cell lines before and after treatment with GSK126 (Fig. 2 and Fig. 5), and GSE49284 for EZH2 mutated cell lines before and after treatment with EPZ6438 (Fig. 5). GSE12195 for normal centroblast, centrocytes and memory B-cells mRNA expression (Suppl. Figure 3).

Field-specific reporting

Please select the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/authors/policies/ReportingSummary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	Analyses of publicly available datasets used all available samples and multiple independent datasets were used to corroborate the reproducibility of the findings. The number of animals used for HPC cells experiments (n > 13) were in accordance with power analyses done based on previous experiment were n = 10 was deemed sufficient.
Data exclusions	Differential expression analysis performed only on cell lines from GSE40972 exhibiting hotspot mutations for EZH2 or EZH2 and UTX wild-type. HT cells were excluded as they harbor a mutation in UTX gene. UTX can in part mimic the effect of EZH2 mutations, thus we couldn't include these cells either in the wild-type or in the EZH2 mutant group. Pfeiffer cells were excluded because they harbor a different mutation in EZH2 (A677).Both of these cell lines were per-excluded from the analysis. In the STORM data analysis, we excluded measurements with extreme numbers of localizations (low or high) which were associated with inflation or deflation of Rg and E values due to technical reasons related to the STORM acquisitions. The detailed procedure is reported in the supplementary note. All measurements are reported in Supplementary Table 7 along with flag indicating whether they were retained or excluded from subsequent analyses.
Replication	Results presented in this study have been confirmed on independent datasets. The replication number for each experiment is reported in each figure legend and individual data
Randomization	The experimental design did not included a randomization of the samples in different group, however experimental observations have been compared to analytically and empirically derived random models to assess their statistical significance as described in the corresponding method sections.
Blinding	For the animal experiments, we measured survival and cell composition. In the initial experimental design we did not include blinding, as we wanted to match cell composition before and after the injection

Reporting for specific materials, systems and methods

Materials & experimental systems		Methods	
n/a	Involved in the study	n/a	Involved in the study
☒	☐ Unique biological materials	☐	☒ ChIP-seq
☐	☒ Antibodies	☐	☒ Flow cytometry
☐	☒ Eukaryotic cell lines	☒	☐ MRI-based neuroimaging
☒	☐ Palaeontology		
☐	☒ Animals and other organisms		
☒	☐ Human research participants		

Antibodies

Antibodies used

Primary Antibodies
Tri-methyl histone H3 (Lys 27) (H3K27me3), Rabbit monoclonal antibody (clone C36B11), Cell Signaling 9733; 1:1000 dilution for western blots, 1:1600 dilution for immunofluorescence, and 1:50 for ChIP

Di-methyl histone H3 (Lys 27) (H3K27me2), Rabbit polyclonal antibody Abcam 194690; 1:500 dilution for western blot

Histone H3, XP Rabbit monoclonal antibody (clone D1H2), Cell Signaling 4499; 1:1000 dilution for western blot

Histone H3 (D2B12) XP® Rabbit mAb (ChIP Formulated), Cell Signaling 4620; 1:50 dilution for ChIP

Normal Rabbit IgG, Rabbit polyclonal antibody, Cell Signaling 2729; 1:500 dilution, or 1ug for ChIP

EZH2, Mouse monoclonal antibody (clone AC22), Cell Signaling 3147; 1:1000 dilution for western blots

β-Actin, Mouse monoclonal antibody (clone 8H10D10), Cell Signaling 3700; 1:1000 dilution for western blot

Ki67, Rabbit monoclonal antibody (clone SP6), Abcam Ab16667; 1:200 dilution

CD3 α-rat APC (clone 17A2), lot 6224663, BD Farmingen 565643; 1:100 dilution

CD45R/B220 α-rat PE-Cy7 (RA3-6B2), lot 6105665, BD Farmingen 552772; 1:100 dilution

Secondary Antibodies
Streptavidin Alexa Fluor-647 conjugate, α-biotin, ThermoFischer S21374 (1:500 dilution)

IRDye 800CW Goat anti-Mouse IgG (H + L), 0.1 mg (1:10000 dilution)

IRDye 680RD Goat anti-Mouse IgG (H + L), 0.1 mg (1:10000 dilution)

Goat anti-Rabbit IgG (H+L) Highly Cross-Adsorbed Secondary Antibody, Alexa Fluor 488, Invitrogen (A-11034) (1:500 dilution)

4',6-Diamidino-2-Phenylindole, Dihydrochloride (DAPI) Invitrogen D1306 (1:5000 dilution)

Validation

1. CD45R/B220 α-PE-Cy7 and CD3 α-APC were used for flow cytometry analyses. Internal negative (unstained) controls were used for all samples.

2. Rabbit α-Ki67 (SP6) was used for histological analysis of tumors. Internal negative control (secondary only) stained slides were used for all samples

3. Tri-methyl histone H3 (Lys 27) (H3K27me3), Rabbit IgG (2729) and Histone H3 (4620) were used for ChIP qPCR/ChIP sequencing experiments. These antibodies were used a part of a kit offered by Cell Signaling Technologies (#9003). The kit contains a positive control Histone H3 Antibody, a negative control Normal Rabbit IgG Antibody and primer sets for PCR detection of the human and mouse ribosomal protein L30 (RPL30) genes. This kit has been extensively used and cited in a number of papers ((1) Orlando, V. (2000) Trends Biochem Sci 25, 99–104., (2) Kuo, M.H. and Allis, C.D. (1999) Methods 19, 425–33., (3) Agalioti, T. et al. (2000) Cell 103, 667–78., (4) Soutoglou, E. and Talianidis, I. (2002) Science 295,1901–4., (5) Mikkelsen, T.S. et al. (2007) Nature 448, 553–60., (6) Lee, T.I. et al. (2006) Cell 125, 301–13., (7) Weinmann, A.S. and Farnham, P.J. (2002) Methods 26, 37–47., (8) Wells, J. and Farnham, P.J. (2002) Methods 26, 48–56.).

4. Tri-methyl histone H3 (Lys 27) (H3K27me3) (9733), Di-methyl histone H3 (Lys 27) (H3K27me2) (194690), Histone H3 (4499), EZH2 (3147) and β-Actin (3700) were used as primary antibodies for western blot analyses. IRDye 800CW Goat anti-Mouse IgG and CD45R/B220 α-rat PE-Cy7 were used as secondary antibodies for western blot analyses.

Histone H3 (4499) has been cited in over 200 peer reviewed journals. Cell Signaling Technology cites no cross reactivity with other histones, and provides a sample western blot for the endogenous detection of Histone H3 in of four separate cells lines (HeLa, C6, COS, and NIH/3T3).

H3K27me2 (194690) has been cited in 2 peer reviewed journals. Abcam provides a sample western blot for the endogenous detection of H3K27me2 in HeLa cells with no detection observed detection of H3K27me2 in E.coli.

H3K27me3 (9733) has been cited in over 200 peer reviewed journals. Cell Signaling Technology states "Tri-Methyl-Histone H3 (Lys27) (C36B11) Rabbit mAb detects endogenous levels of histone H3 only when tri-methylated on Lys27. The antibody does not cross-react with non-methylated, mono-methylated or di-methylated Lys27. In addition, the antibody does not cross-react with mono-methylated, di-methylated or tri-methylated histone H3 at Lys4, Lys9, Lys36 or Histone H4 at Lys20." Cell signaling provides a sample western blot for the endogenous detection of H3K27me3 in four separate cell lines (HeLa, C6, COS, and NIH/3T3).

EZH2 (3147) has been cited in over 70 peer reviewed journals. Cell signaling provides a sample western blot for the endogenous detection of EZH2 in four separate cell lines (T47D, MCF7, SEM, MDA-MB134).

β-Actin (3700) has been cited in over 650 peer reviewed journals. Cell signaling provides a sample western blot for the endogenous detection of β-Actin in five separate cell lines (COS, HeLa, C2C12, C6, CHO).

5. Tri-methyl histone H3 (Lys 27) (H3K27me3) (9733) was used as a primary antibody for immunofluorescence staining (1:1600). Goat anti-Rabbit IgG Alexa Fluor 488 (1:500) (A-11034) and DAPI (D1306) (D1306) (1:5000) were used as secondary antibodies for immunofluorescence stain.

Internal positive and negative (secondary only) control stained slides were used for validation. In addition, Cell Signaling cites over 200 peer reviewed journals where H3K27me3 (9733) has been used, and states that "Tri-Methyl-Histone H3 (Lys27) (C36B11) Rabbit mAb detects endogenous levels of histone H3 only when tri-methylated on Lys27. The antibody does not cross-react with non-methylated, mono-methylated or di-methylated Lys27. In addition, the antibody does not cross-react with mono-methylated, di-methylated or tri-methylated histone H3 at Lys4, Lys9, Lys36 or Histone H4 at Lys20." Cell signaling provides an example image of HeLa cells immunolabeled with the (H3K27me3) (9733) antibody. Goat anti-Rabbit IgG Alexa Fluor 488 (A-11034) has been cited in over 500 peer reviewed journals, and uses a process of cross-absorption to remove antibodies that potentially cross react with other species, which increases specificity and reduces background.

6. Streptavidin α-biotin Alexa Fluor-647 (1:500) (S21374) and custom designed oligo probes (MYcroarray) were used for DNA FISH and STORM imaging experiments. Internal negative control (secondary only) stained slides were used for validation of the secondary antibody.

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)	DoHH2 (ACC-47) DSMZ, WSU-DLCL2 (ACC-575) DSMZ, SU-DHL-10 (ACC-576) DSMZ, OCI-LY19 (ACC-528) DSMZ, Karpas-422 and Toledo and FI-5-12 cells were obtained from Wendel lab (MSKCC)
Authentication	All human cell lines were authenticated by short tandem repeat (STR) profiling, performed by Microsynth, CH.
Mycoplasma contamination	cells obtained for Wendel lab were tested and were negative for mycoplasma contamination, cell bought from DSMZ have not been tested for mycoplasma contamination.
Commonly misidentified lines (See ICLAC register)	No commonly misidentified cell lines were used.

Animals and other organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research

Laboratory animals	mus musculus, C57BL6, females, 8-10 weeks old, HPC were isolated from embryo at 13.5-14.5 days.
Wild animals	This study did not involve wild animals
Field-collected samples	This study did not involve field-collected samples

ChIP-seq

Data deposition

☒ Confirm that both raw and final processed data have been deposited in a public database such as GEO.

☒ Confirm that you have deposited or provided access to graph files (e.g. BED files) for the called peaks.

Data access links	H3K27me3 for OCI-Ly7, DOHH-2 and Karpas-422 were downloaded from ENCODE

Data access links *May remain private before publication.*	H3K27me3 for WSU-DLCL2 were downloaded from GSE40970 H3K4me3 for OCI-Ly7 and Karpas-422 were downloaded from ENCODE H3K27me3 and RNAseq for OCI-Ly19-WT and OCI-Ly19-EZH2 Y646F were generated as described in the manuscript and deposited in GSE114270, the token for reviewer access is admhsguudlkldwv HiC matrices that can be visualized with Juicebox have been deposited in Zenodo in a private link https://zenodo.org/record/1244183? token=eyJhbGciOiJIUzI1NiIsImV4cCI6MTUyODU4MTU5OSwiaWF0IjoxNTI1OTQ1Mjg0fQ.eyJkYXRhIjp7InJlY2lkIjoxMjQ0MTgzfSwiaWQiOjUzOSwicm5kIjoiYTJhNTIzMjIifQ.3lznpUuQFO-Nkbu895VkhvmUhNieYFFsZ0ZCqedFuHw
Files in database submission	FASTQ and .bam alignment files for H3K27me3 and input (OCI-Ly19-WT and OCI-Ly19-Y646F) .bed peak lists for H3K27me3 (OCI-Ly19-WT and OCI-Ly19-Y646F)
Genome browser session (e.g. UCSC)	N/A

Methodology

Replicates	OCI-LY19-WT and OCI-LY19-Y646F: 3 replicates per condition
Sequencing depth	OCI-LY19-WT: 149102952 mapped (MAPQ>0) reads for H3K27me3 ChIP and 48538036 for the input OCI-LY19-Y646F: 122262552 mapped (MAPQ>0) reads for H3K27me3 ChIP and 46698494 for the input
Antibodies	H3K27me3: Cell Signaling 9733
Peak calling parameters	Peaks were called with SICER_V1.1, using hg19, fragment size 250bp; effective genome size fraction 0.86; window size 750 bp; gap size 3; redundancy threshold 1; FDR 0.001
Data quality	- Mapping quality filtering to exclude secondary alignments and reads aligning to multiple locations (MAPQ>0) - Strict FDR threshold to retain SICER peaks (FDR<0.001) - Visual inspection of ChIP, input and peak tracks on IGV
Software	- Bowtie2-2.2.9 for the alignment - Samtools-1.6 for removing PCR duplicates, sorting and filtering - SICERpy (wrapper for SICER_V1.1) for peak calling - bedGraphToBigWig v 4 was used to convert bedGraph to bigWig file - IGV genome browser for visualization - Custom scripts under https://github.com/CSOgroup/Donaldson-et-al-scripts for downstream analyses

Flow Cytometry

Plots

Confirm that:

☒ The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).

☒ The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).

☒ All plots are contour plots with outliers or pseudocolor plots.

☒ A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation	Cells were retrovirally transduced with either shRNA for Sesn1, Foxo3, or MSCV-IRES-GFP or MSCV-IRES-DsRed vectors. Cells were acquired in the indicated days
Instrument	The number of fluorescent cells were monitored by FACS (Guava Millipore EasyCyte HT System or LSRII Snoopy , Becton Dickinson).
Software	The data were analyzed using FlowJo.
Cell population abundance	All cells were considered in the analyses.
Gating strategy	In the initial gating FSC/SCC we only excluded debris or dead cells that in general accounted for less than 10% of the population, all living cells were considered in the analyses.

☐ Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.

Appendix 2: CV

Maria Donaldson Collier
Maria.Donaldson@epfl.ch
+41.78.834.2883
Research Unique Identifier: 0000-0001-6403-6713

EDUCATION

École Polytechnique Fédérale de Lausanne August 2015 – Present
Lausanne, Switzerland
Doctoral Student in the molecular life science program at the Swiss Institute for Experimental Cancer Research

University of Vermont 2007 – 2011
Burlington, Vermont, U.S.A.
Bachelor of Arts degree in Biology, minor in Chemistry

Sea Education Association (SEA) 2009
Woods Hole, Massachusetts, U.S.A.
Marine biodiversity and conservation semester at sea program

Foxcroft School 2003 – 2007
Middleburg, Virginia, U.S.A.

RESEARCH EXPERIENCE

Doctoral Student, The Swiss Institute for Experimental Cancer Research August 2015 – Present
École Polytechnique Fédérale de Lausanne, Lausanne, Switzerland
Mentored by Dr. Elisa Oricchio

- Examining how EZH2 oncogenic mutations drive transcriptional and structural changes within topologically associating domains (Donaldson et al., *Nature Genetics*; Oricchio et al., *Science Translational Medicine*; Donaldson et al., *Molecular and Cellular Oncology*)
  Specialized skills established : Hi-C, UMI-4C and ChIP sequencing, DNA in situ hybridization, preparation of samples for next generation RNA sequencing and DNA methylation profiling
- Development of a robust in vivo model for studying central nervous system primitive neural ectodermal tumors (CNS-PNETs) (Manuscript in preparation)
  Specialized skills established : stereotaxic injections (murine), human stem cell culture and differentiation, confocal imaging, and laser capture microdisection microscopy
- Deciphering the metabolic and signaling dependencies of CNS-PNETs to interrogate novel detection and therapeutic techniques (Manuscript in preparation)

Senior Research Technician, The Center for Cancer Research June 2012 – July 2015
Massachusetts General Hospital, Harvard Medical School, Boston, MA, USA
Mentored by Dr. Daniel Haber, Dr. Shyamala Maheswaran, Dr. Nicola Aceto, and Dr. Min Yu

Exploring the role of circulating tumor cells (CTCs) in breast cancer.

- Development of a microfluidic device to capture CTC clusters (Sarioglu et al., *Nature Methods*, April 2015)
- Isolation and characterization of circulating tumor cell clusters (CTC clusters) in cancer disease progression and metastasis (Aceto et al., *Cell*, August 2014; Aceto et al., *Molecular Cancer Research* 2018)
- Establishing patient-derived breast cancer CTC-cultures for individualized treatment through drug susceptibility testing and monitoring for acquired mutations (Yu et al., *Science*, July 2014)
- Characterizing the role of epithelial to mesenchymal transition (EMT) in CTCs isolated from patients with metastatic breast cancer (Yu et al., *Science*, February 2013)
  Specialized skills established : CTC isolation, CTC culture (including organoid culture), single cell microscopic manipulation, RNA-in situ hybridization, tail vein injection, mammary gland injection, cardiac puncture for whole blood collection and *in vivo* bioluminescent imaging (BLI) of live mice.

Research Intern August 2011 - December 2011
Bermuda Institute of Ocean Sciences, St. George's, Bermuda; Mentored by Dr. Andreas Andersson

- Examining the effects of ocean acidification on $CaCO_2$ precipitation and carbonate shell formation in Bermuda native coral reef species
  Specialized skills established : certified AAUS scientific scuba diver, PADI-certified advanced open-water and rescue diver, and underwater photography

Research Assistant, Department of Ecology May 2010 - May 2011
University of Vermont, Burlington, Vermont; Mentored by Dr. Edmund Hart and Dr. Nicholas Gotelli

- Studying the effects of climate change on the density-dependent population dynamics of aquatic invertebrates.

PUBLICATIONS

1. EZH2 oncogenic mutations drive transcriptional and structural changes within topologically associating domains. *Nature Genetics*, January 2019. https://doi.org/10.1038/s41588-018-0338-y.
Maria Donaldson Collier*, Stephanie Sungalee*, Marie Zufferey*, Daniele Tavernari*, Natalya Katanayeva, Marco Mina, Kyle M. Douglass, Timo Rey, Franck Reynaud, Suliana Manley, Giovanni Ciriello, and Elisa Oricchio.

2. AR expression in breast cancer CTCs associates with bone metastases.
Molcular Cancer Research, April 2018, 16, no. 4 (April 2018): 720–27. https://doi.org/10.1158/1541-7786.MCR-17-0480.
Nicola Aceto, Aditya Bardia, Ben S. Wittner, **Maria C. Donaldson**, Ryan O'Keefe, Amanda Engstrom, Francesca Bersani, Yu Zheng, Valentine Comaills, Kira Niederhoffer, Huili Zhu, Olivia Mackenzie, Toshi Shioda, Dennis Sgroi, Ravi Kapur, David T. Ting, Beverly Moy, Sridhar Ramaswamy, Mehmet Toner, Daniel A. Haber, and Shyamala Maheswaran.

3. Pan-SRC kinase inhibition blocks B-cell receptor oncogenic signaling in non-Hodgkin lymphoma.
Blood 131, no. 21 (May 24, 2018): 2345–56. https://doi.org/10.1182/blood-2017-10-809210.
Elena Battistello, Natalya Katanayeva, Elie Dheilly, Daniele Tavernari, **Maria C. Donaldson**, Luca Bonsignore, Margot Thome-Miazza, Amanda L. Christie, Mark A. Murkami, Olivier Michielin, Giovanni Ciriello, Vincent Zoete, and Elisa Oricchio.

4. Sestrin1, a tumor suppressor than can be rescued.
Molecular & Cellular Oncology 4, no. 6 (November 2, 2017): e1365107. https://doi.org/10.1080/23723556.2017.1365107.
Maria C. Donaldson, Natalya Katanayeva, and Elisa Oricchio.

5. Genetic and epigenetic inactivation of SESTRIN1 controls mTORC1 and response to EZH2 inhibition in follicular lymphoma.
Science Translational Medicine 9, no. 396 (June 28, 2017) https://doi.org/10.1126/scitranslmed.aak9969.
Elisa Oricchio, Natalya Katanayeva, **Maria C. Donaldson**, Stephanie Sungalee, Passion P. Joyce, Wendy Béguelin, Elena Battistello, Viraj R. Sanghvi, Man Jiang, Yanwen Jiang, Matt Teater, Anita Parmigiani, Andrei V. Budanov, Fong Chun Chan, Sohrab P. Shah, Robert Kridel, Ari M. Melnick, Giovanni Ciriello, Hans-Guido Wendel.

6. A microfluidic device for label-free, physical capture of circulating tumor cell-clusters.
Nature Methods 12, no. 7 (July 2015): 685–91. https://doi.org/10.1038/nmeth.3404.
A. Fatih Sarioglu, Nicola Aceto, Nikola Kojic, **Maria C. Donaldson**, Mahnaz Zeinali, Bashar Hamza, Amanda Engstrom, Huili Zhu, Tilak K. Sundaresan, David T. Miyamoto, Xi Luo, Aditya Bardia, Ben S. Wittner, Sridhar Ramaswamy, Toshi Shioda, David T. Ting, Shannon L. Stott, Ravi Kapur, Shyamala Maheswaran, Daniel A. Haber, and Mehmet Toner.

7. Circulating tumor cell clusters are oligoclonal precursors of breast cancer metastasis.
Cell, August 2014 158, no. 5 (August 2014): 1110–22. https://doi.org/10.1016/j.cell.2014.07.013.
Nicola Aceto, Aditya Bardia, David T. Miyamoto, **Maria C. Donaldson**, Ben S. Wittner, Joel A. Spencer, Min Yu, Adam Pely, Amanda Engstrom, Huili Zhu, Brian W. Brannigan, Ravi Kapur, Shannon L. Stott, Toshi Shioda, Sridhar Ramaswamy, David T. Ting, Charles P. Lin, Mehmet Toner, Daniel A. Haber, and Shyamala Maheswaran.

8. Ex vivo culture of circulating breast tumor cells reveals individualized drug susceptibility.
Science 345, no. 6193 (July 11, 2014): 216–20. https://doi.org/10.1126/science.1253533.
Min Yu, Aditya Bardia, Nicola Aceto, Francesca Bersani, Marissa W. Madden, **Maria C. Donaldson**, Rushil Desai, Huili Zhu, Valentine Comaills, Zongli Zheng, Ben S. Wittner, Petar Stojanov, Elena Brachtel, Dennis Sgroi, Ravi Kapur, Toshishiro Shioda, David T. Ting, Sridhar Ramaswamy, Gad Getz, A. John Iafrate, Cyril Benes, Mehmet Toner, Shyamala Maheswaran, and Daniel A. Haber.

9. Circulating breast tumor cells exhibit dynamic changes in epithelial and mesenchymal composition.
Science 339, no. 6119 (February 1, 2013): 580–84. https://doi.org/10.1126/science.1228522.
Min Yu, Aditya Bardia, Ben S. Wittner, Shannon L. Stott, M.E. Smas, David T. Ting, Steven J. Isakoff, Jordan C. Ciciliano, Marissa N. Wells, Ajay M. Shah, Kyle F. Concannon, **Maria C. Donaldson**, Lecia V. Sequist, Elena Brachtel, Dennis Sgroi, Jose Baselga, Sridhar Ramaswamy, Mehmet Toner, Daniel A. Haber, and Shyamala Maheswaran.

INVITED SPEAKER (Selected)

"EZH2 oncogenic mutations drive epigenetic, transcriptional, and structural changes within topologically associating domains"
Invited talk, University of Lausanne department of Biochemistry, Lausanne, Switzerland. October 2018.
Invited talk, Netherlands Cancer Institute, Amsterdam, Netherlands. January 2019
Invited talk, Institute of Cancer Research, London, UK. January 2019
Invited talk, Hubrecht Institute, Utrecht, Netherlands. March 2019

"EZH2 oncogenic mutations alter structure and function of topologically associating domains."
Poster prize, Federation of American Societies for Experimental Biology (FASEB) Hematologic Malignancy conference. Vermont, USA. July 2017.

"Deciphering the metabolic and signaling dependencies of central nervous system primitive neural ectodermal tumors (CNS-PNET) to interrogate novel therapeutic and detection techniques."
Selected Talk, LIMNA Symposium. Lausanne, Switzerland. May 2017.

"Identifying the Role of Circulating Tumor Cell Clusters in Breast Cancer Metastasis."
PhD interview; invited talk. École Polytechnique Fédérale de Lausanne, Switzerland. January 2015.

"Ex vivo culture of circulating breast tumor cells reveals individualized drug susceptibility."
Poster prize, Massachusetts General Hospital, Harvard Medical School, USA. October 2014.

GRANTS

Sydney L. Wright Fellowship Grant-In-Aid, awarded by the Bermuda Institute of Ocean Sciences, Bermuda.

TEACHING

2016-2017 (2 semesters) Life science integrated laboratory I and II - supervised students in applying basic molecular biology techniques in the laboratory

2016-2018 (4 semesters) Cellular and Molecular Biology I and II – supervised exercise courses to apply the basic principles of molecular and cellular biology addressed during the corresponding lecture each week

Spring 2017 (1 semester) - supervised a masters student (MS) project during laboratory immersion

Ce document a été imprimé au Centre d'impression EPFL,
imprimerie climatiquement neutre, certifiée myClimate.