Modeling PRC2 Alterations in Cancer, a Window Toward a Better Comprehension of EZH2 Mutation in Follicular Lymphoma

Pierre Romero

université
PARIS-SACLAY

Modeling PRC2 alterations in cancer, a window toward a better comprehension of *EZH2* mutation in follicular lymphoma

Modélisations des altérations de PRC2 dans les cancers, vers une meilleure compréhension du rôle de la mutation d'EZH2 dans les lymphomes folliculaires

Thèse de doctorat de l'université Paris-Saclay

École doctorale n° 582, Cancérologie : Biologie - Médecine - Santé (CBMS)
Spécialité de doctorat : Sciences de la vie et de la santé
Unité de recherche : Institut Curie, PSL Research University, CNRS UMR 3215, INSERM U934, 75248 Paris Cedex 05, France
Référent : Faculté de médecine

Thèse présentée en visioconférence totale à Paris-Saclay le 14 Octobre 2021, par

Pierre ROMERO

Composition du Jury

Franck MORSCHHAUSER PU-PH, Université de Lille, CHRU de Lille	Président
Sandrine ROULLAND Chargée de recherche INSERM, Université d'Aix-Marseille, CIML	Rapporteur & Examinatrice
Adrian BRACKEN Associate Professor, Trinity College at University of Dublin, Smurfit Institute of Genetics	Rapporteur & Examinateur
Alexandra TRAVERSE-GLEHEN PU-PH, Université Lyon 1, HCL	Examinatrice

Direction de la thèse

Anne VINCENT-SALOMON HDR, Université Paris-Saclay, Institut Curie	Directrice de thèse
Raphaël MARGUERON Directeur de recherche 1 INSERM, Sorbonne Université, Institut Curie	Co-Directeur de thèse

Thèse de doctorat

NNT : 2021UPASL073

Modeling PRC2 alterations in cancer, a window toward a better comprehension of *EZH2* mutation in follicular lymphoma

I am deeply grateful to my reviewers and jury members, Drs. Sandrine Roulland & Adrian Bracken, for their time in evaluating my manuscript and valuable advice to allow me to improve it.

I am honored and grateful to the jury members, Prs. Alexandra Traverse-Glehen & Franck Morschhauser, for accepting to judge the present work.

I warmly thank the members of my thesis advisory committee, Dr. Sophie Postel-Vinay & Pr. Thierry Molina, for their goodwill and constructive feedback during my two TAC meetings.

I thank my co-director, Dr. Anne Vincent-Salomon, for introducing me to Raphaël and contributing in the funding of the 4$^{th.}$ year of this thesis. I thank my co-director and supervisor, Dr. Raphaël Margueron, for his daily guidance and to whom I am indebted. I thank the training unit of Institut Curie for allowing me to experience science.

Seynabou Diop & Daniel Ming-Kang Lee, you made this journey worth it. And as the journey ends, I also wish to address my gratitude to Laia Richart-Gines, who epitomizes a scientific mind to me and tamed (but not killed) my rebellion to computational science, to Katia Ancelin, who read and helped me tweak my manuscript and to Roberta Ragazzini, who taught me how to ChIP real slow. I thank all the members of the Margueron lab for being themselves, a PhD is a human experience after all.

I thank Julien Masliah-Planchon and Corinne Ajrizov as well as their colleagues from the unit of genetics, the members of the Center of Biological Resources, especially Céline Méaudre and Aurore Loistron, Laetitia Fuhrmann from the department of pathology, the members of ICGex platform and NGS lab, and Frédérique Khunowski for their significant contribution and support during this project.

To my silent yet powerful army, thank you.

« Dans l'existence, surviennent des heures, des détails qui sembleraient n'avoir qu'une valeur de dernier ordre et qui se gravent minutieusement dans la mémoire, tandis que d'autres, mille fois plus importants, n'y laissent aucune trace. » *Loti, Le roman d’un enfant*

Next generation sequencing unveiled the pivotal role in cancer of general mechanisms involved in the control of gene expression including chromatin modifiers. The Polycomb Repressive Complex 2 “PRC2” catalyzes the methylation of H3K27, associated with inactive genes. PRC2 is found to be altered in various cancer types and the consequences of these alterations depend on both tumor type and the nature of the alteration. Some of them result in abrogation of the deposition of H3K27me3, whereas other, namely the mutation Y646 of the catalytic sub-unit of PRC2 *EZH2* present in up to 25% of follicular lymphoma (FL), leads to both an increase and a redistribution of H3K27me3. Thus, PRC2 can be alternatively considered as a tumor suppressor or as an oncogenic protein complex given the context. This PhD thesis aimed at clarifying the mechanistic consequences of these alterations through two complementary approaches.

1/ Mechanistic approach: Modeling main alterations of PRC2 in cancers in an isogenic cell line model

We used an immortalized murine embryonic fibroblast (mEF) cell line harboring a conditional deletion of *Ezh2* and recapitulating *Eed*-KO, H3.3K27M, *Ezh2*$^{Y641F/WT}$ and *Ezh2*$^{Y641F/-}$+/- *Ezh1*-KO genotypes. Loss-of-function of PRC2 leads to a decrease of H3K27me3 and a reactivation of PRC2 target genes. However, we don’t observe generation of new target genes nor a gain of the mark, especially upon H3.3K27M mutation. Conversely, Ezh2 Y641F induces an increase of H3K27me3, regardless of the presence of either an *Ezh2* WT allele or *Ezh1* to ensure monomethylation. Our results suggest that co-factors of PRC2 restore the capacity of PRC2-Y641 to catalyze monomethylation. Ezh2 Y641F leads to a redistribution of H3K27me3 in *Ezh2*$^{Y641/WT}$ cells with a flattening of the peaks and a global propagation of the deposition. Such property seems to be the direct consequence of hyperactivity of PRC2-Y641 since partial inhibition of its activity restores a more “sawtooth” pattern. Los- of-function alterations of PRC2 are associated with a global increase of H3K27ac yet with a relative decrease of the height of the peaks, while Ezh2 Y641F induces an increase of H3K27ac at the whole genome level. Transcriptional response to EZH2 inhibitors is quantitatively similar between WT and *Ezh2*$^{Y641F/WT}$ cells, but the nature of responsive genes is different: *Ezh2*$^{Y641F/WT}$ cells unexpectedly re express genes involved in antigenic presentation, in a similar way as observed in some B cell lymphoma models.

2/ Translational approach: **Constitution and follow-up of a cohort of follicular lymphoma cases WT and mutant for *EZH2***

Sequencing of exons 16/18 of *EZH2* from DNA extracted out of 160 FL samples allowed the identification of 18.8% mutated cases. In a restricted cohort of 32 cases (21 mutant *EZH2*^{Y646}, 11 WT) with a median follow-up of 11.5 years, we conducted an integrative study including coupled sequencing and CNV analysis of 571 genes involved in cancer, RNA-Seq and H3K27me3 ChIP-Seq. For 13 cases, ChIP-Seq and RNA-Seq have been performed in ≥2 longitudinal time-point samples/patient. We confirmed the high frequency of *KMT2D* and *CREBBP* mutations without specific co-occurrence with *EZH2* mutations. Transcriptomic signature of *EZH2*^{Y646} cases showed a similar number of up and down regulated genes, with an enrichment in pathways involved in microenvironmental crosstalk for the repressed genes. EZH2 Y646 induces a redistribution of H3K27me3 along with an enrichment at the gene bodies and a depletion at the promoter regions. Enrichment at the gene bodies is more marked on the downregulated genes in the mutant background, whereas upregulated genes do not show a noticeably different H3K27me3 pattern according to the genotype. Finally, B cell-specific enhancers are enriched in H3K27me3 in *EZH2* mutant cases, along with differential transcriptional repression compared with WT cases.

Le séquençage à haut débit a permis d'identifier dans les cancers des mécanismes généraux de contrôle de l'expression des gènes dont les modificateurs de la chromatine. Le complexe de répression transcriptionnelle Polycomb PRC2 catalyse la méthylation de H3K27, associée aux gènes inactifs. PRC2 est l'objet de diverses altérations dans les cancers dont les conséquences sur l'activation des gènes varient selon le type tumoral et la nature de l'altération. Certaines de ces altérations résultent en une abrogation de H3K27me3, alors que d'autres, comme la mutation Y646 de la sous-unité catalytique de PRC2 *EZH2* rencontrée dans 25% des lymphomes folliculaires (LF), induit une augmentation et une redistribution de H3K27me3. PRC2 peut donc être considéré comme suppresseur de tumeur ou oncogène selon le contexte. Cette thèse a eu pour objectif de clarifier les conséquences mécanistiques de ces altérations au travers de deux approches complémentaires.

1/ Approche mécanistique : Modélisation des altérations de PRC2 dans les cancers dans un modèle cellulaire isogénique

Nous avons utilisé une lignée de fibroblastes embryonnaires murins immortalisés (mEF) porteurs d'une délétion conditionnelle d'*Ezh2* et récapitulant les génotypes *Eed*-KO, H3.3K27M, *Ezh2*$^{Y641F/WT}$ et *Ezh2*$^{Y641F/-}$ +/- *Ezh1*-KO. La perte de fonction de PRC2 conduit à une diminution d'H3K27me3 et la réactivation de cibles de PRC2. Toutefois, nous n'observons pas la génération de nouvelles cibles, ni le gain de cette marque, liés à H3.3K27M notamment. A l'inverse, Ezh2 Y641F induit une augmentation de H3K27me3 et ne nécessite pas la présence d'un allèle WT ou d'*Ezh1* pour assurer la monométhylation. Nos résultats suggèrent l'implication des cofacteurs de PRC2 dans la capacité de PRC2-Y641 à catalyser la monométhylation. Ezh2 Y641F entraîne une redistribution de H3K27me3 dans les cellules *Ezh2*$^{Y641F/WT}$ avec un aplatissement des pics mais une propagation globale de la déposition. Cette propriété semble être une conséquence directe de l'hyperactivité de PRC2-Y641F car l'inhibition partielle de son activité restaure un profil plus « vallonné ». Les pertes de fonction de PRC2 sont associées à une augmentation globale de H3K27ac avec une diminution relative au niveau des pics alors que Ezh2 Y641F induit une augmentation globale de H3K27ac sur l'ensemble du génome. La réponse transcriptionnelle aux inhibiteurs d'EZH2 est quantitativement similaire entre les lignées WT et *Ezh2*$^{Y641F/WT}$ mais la nature des gènes est différente : la lignée *Ezh2*$^{Y641F/WT}$ montre de façon inattendue une réexpression de gènes impliqués dans la présentation antigénique, comme observé dans des modèles de lymphomes B.

2/ Approche translationnelle : Constitution d'une cohorte de lymphomes folliculaires *EZH2* WT et mutés

Le séquençage des exons 16/18 d'*EZH2* à partir de 160 cas de LF a permis d'identifier 18.8% de cas mutés. Sur une cohorte restreinte de 32 patients (21 cas *EZH2*^{Y646}, 11 *EZH2* WT) avec un suivi clinique médian de 11.5 ans, nous avons réalisé, sur au moins un temps biopsique, une étude intégrative associant séquençage et CNV de 571 gènes impliqués dans le cancer, RNA-Seq et ChIP-Seq H3K27me3. Pour 13 cas, ChIP-Seq et RNA-Seq ont été réalisés de façon séquentielle sur ≥2 biopsies/patient. Nous confirmons la prévalence des mutations de *KMT2D* et *CREBBP* sans co-occurrence spécifique avec les mutations d'*EZH2*. La signature transcriptionnelle des cas mutés montre autant de gènes activés que réprimés, avec un enrichissement dans les voies du microenvironnement pour les gènes réprimés. La mutation induit une redistribution de H3K27me3 avec un enrichissement sur les corps des gènes et une diminution sur les régions promotrices. L'enrichissement sur les corps des gènes est plus marqué pour les gènes réprimés par EZH2 Y646, les gènes activés ne présentent pas de profil différent selon le génotype. Enfin, les enhancers spécifiques aux cellules B sont enrichis dans les cas mutés, en lien avec une répression transcriptionnelle différentielle.

INTRODUCTION

RESULTS

A

ABC : Activated B cell
ACVR1 : Activin A receptor type 1
AEBP2 : AE binding protein 2
AID : Activation-induced cytidine deaminase
AML : Acute myeloid leukemia
ARID1A : AT-rich interaction domain 1A
ASXL1 : ASXL transcriptional regulator 1

B

B2M : Beta-2-Microglobulin
BAD : BCL2 associated agonist of cell death
BAK : BCL2 Antagonist/Killer 1
BAX : BCL2–associated X
BCAT1 : Branched chain amino acid transaminase 1
BCL2 : B cell lymphoma 2
BCL7A : BAF Chromatin Remodeling Complex Subunit
BCR : B cell receptor
BET : Bromodomain and extra-terminal motif
ß-TrCP : Beta-transducin repeat-containing proteins
BID : Bis in die/ twice per day
BIM : BCL2-like protein 11
Blimp1 : B-lymphocyte-induced maturation protein 1
BMI1 : B cell-specific Moloney murine leukemia virus integration site 1
bp : base pair
BRD : Bromodomain
BTLA : B and T lymphocyte associated

C

CARD11 : Caspase recruitment domain family member 11
CART-T : Chimeric antigen receptor T-cells
CATACOMB : Catalytic antagonist of Polycomb
CBX : Chromobox homolog
CDKN : Cyclin dependent kinase inhibitor
ChIP-Seq : Chromatin immunoprecipitation-sequencing
CHOP/R-CHOP/R-CVP : Cyclophosphamide hydroxorubicin vincristine and prednisone/rituximab - CHOP/R-cyclophosphamide vincristine and prednisone
CMG : Chromatin modifying gene
CML : Chronic myeloid leukemia
COMPASS : Complex proteins associated with Set1
CPC : Common precursor cell
CpG : Cytosine-phosphate-guanine
CREBBP : CREB Binding Protein
CRISPR-Cas9 : Clustered regularly interspaced short palindromic repeats/CRISPR associated protein 9
CT : Computerized tomography
CUT&RUN: Cleavage under target & release using nuclease
CXCL12 : C-X-C motif chemokine ligand 12

D

DC-SIGN : DC-specific intercellular adhesion molecule-3 grabbing nonintegrin
DIPG : Diffuse intrinsic pontine glioma
DNA : Deoxyribonucleic acid
DNMT : DNA methyl transferase
DZ : Dark zone
DZNep : 3-Deazaneplanocin A
DLBCL : Diffuse large B cell lymphoma

E

E2F-pRB : E2F-retinoblastoma tumor suppressor protein
EED : Embryonic ectoderm development
EIP : Earliest inferable precursor
EP300 : E1A binding protein p300
EPOP : Elongin BC and polycomb repressive complex 2 associated protein
EZH1/2 : Enhancer of zeste homolog 1/2
EZHIP : EZH1/2 inhibitory protein

F

FBXL10 : F-box and leucine-rich repeat 11
FDC : Follicular dendritic cell
FDG-PET : Fluorodeoxyglucose-positron emission tomography
FFPE : Formalin-fixed paraffin-embedded
FL/tFL : Follicular lymphoma/transformed FL
FLIPI : FL international prognostic index
FLLC : Follicular lymphoma-like cells
FOXO1 : Forkhead box O protein O1

G

GC : Germinal center
GCB: Germinal center B cell

GELA : Groupe d'étude des lymphomes de l'adulte
GWAS : Genome-wide association study

H

H2AUb : Histone H2A ubiquitination
H3K27 : Histone H3 lysine 27
H3K36 : Histone H3 lysine 36
H3K4 : Histone H3 lysine 4
H3K9 : Histone H3 lysine 9
HDAC : Histone deacetylase
HE : Hematoxylin & eosin
HIST1HE1 : H1.4 linker histone
HLA : Human leukocyte antigen
HMT : Histone methyl transferase
HSC : Hematopoietic stem cell
HSP : Heat shock protein

I

IgH : Immunoglobulin heavy chain
IHC : Immunohistochemistry
IKZF1 : IKAROS family zinc finger 1
IL : Interleukin
INI1 : Integrase Interactor 1
IRF4 : Interferon regulatory factor 4
ISFN : In situ follicular neoplasia

J

JAK/STAT : Janus kinase/signal transducer and activator of transcription
JARID2 : Jumonji and AT-rich interaction domain containing 2
JMJD3 : Jumonji domain-containing protein D3

K

KAT : Lysine acetyltransferase
KDM : Lysine demethylase
KMT2D/C : Lysine methyltransferase 2D/C

L

L3MBTL2 : L3MBTL histone methyl-lysine binding protein 2
LINC : LIN complex
lncRNA : long non coding RNA
LOH/cnLOH : Loss of heterozygosity/copy number neutral LOH
LSC : Leukemia stem cell
LZ : Light zone

M

MACS2 : Model-based Analysis of ChIP-Seq
MAX-MGA : MYC associated factor X - MAX gene associated protein
MDS : Myelodysplastic syndrome
ME : Microenvironment
MEF : Mouse embryonic fibroblast
MEF2B : Myocyte enhancer factor 2B
MEK : MAP/ERK kinase
mESC : Mouse embryonic stem cell
MHC : Major histocompatibility complex
MLL : Mixed-lineage leukemia
MPN : Myeloproliferative neoplasm
MPNST : Malignant peripheral nerve sheath tumor
MTF2 : Metal-response element-binding transcription factor 2
mTORC1 : mammalian Target of rapamycin complex 1
MYC : V-Myc avian myelocytomatosis viral oncogene Homolog
MYD88 : Myeloid differentiation primary response 88

N

NF-1 : Neurofibromatosis 1
NF-κB : Nuclear factor-kappa B
NGS : Next generation sequencing
NO66 : Nucleolar protein 66
NOTCH1 : Notch homolog 1, translocation-associated
NOTCH2 : Neurogenic locus notch homolog protein 2
NSD2 : nuclear receptor binding SET domain protein 2
NURD : Nucleosome remodeling and deacetylase
NURF : Nucleosome remodeling factor

O

OS : Overall survival

P

PALI1/2 : PRC2 associated LCOR isoform 1/2
PcG : Polycomb group
PCGF : Polycomb group ring finger
PCL : Polycomb-like homolog
PDX : Patient derived xenograft
PFA : Posterior fossa group A

PFS : Progression free survival
PHC : Polyhomeotic homolog
PHF : PHD finger protein
Pho-RC : Pho-Repressive complex
PI3K : Phosphoinositide 3-kinase
PIFL : Partial involvement by follicular lymphoma
PIP : PALI interaction with PRC2
PMBCL : Primary-mediastinal large B cell lymphoma
POD24 : Progression of disease within 2 years
POLII : RNA polymerase II
PRC1&2 : Polycomb repressive complex 1&2
PRDM1 : PR/SET domain 1
PR-DUB : Polycomb repressive deubiquitinase
PRIMA : Primary rituximab and maintenance
PTM : Post translational modification
PUMA : BCL2 Binding Component 3

Q

qPCR : quantitative Polymerase chain reaction

R

Ras : Rat sarcoma virus
RBAP46/48 : RB-associated protein 46/48
RING1A/B : Really interesting new gene 1 protein A/B
RNA : Ribonucleic acid
RNA-Seq : RNA - sequencing
RRAGC : Ras Related GTP Binding C
RUNX1 : RUNX family transcription factor 1
RYBP : RING1 and YY1 binding protein

S

SAH : S-adenosyl-L-homocysteine hydrolase
SAM : S-adenosyl-l-methionine
SANT/preSANT : Swi3, Ada2, N-CoR, and TFIIIB/pre SANT
SCM : Scm homolog
SESTRIN1 : P53 regulated PA26 nuclear protein
SET : Su(var)3-9, enhancer-of-zeste and trithorax
SHM : Somatic hypermutation
SIN3 : Scaffolding factor SIN3
SMARCB1 : SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily B, member 1
SNP : Single nucleotide polymorphism
SNV : Single nucleotide variant
SOCS1 : Suppressor of cytokine signaling protein 1
STAT5 : Signal transducer and activator of transcription 5
SUZ12 : Suppressor of zeste 12
SWI/SNF : SWItch/sucrose non-fermentable

T

TAD : Topologically associated domain
T-ALL : T-cell acute lymphoblastic leukemia
TAM : 4--Hydroxytamoxifen
TCR : T-cell receptor
TET2 : Ten-eleven translocation 2
TF : Transcription factor
TID : Ter in die/three times per day
TKI : Tyrosine kinase inhibitor
TMTV : Total metabolic tumor volume
TNFRSF14/HVEM : TNF Receptor Superfamily Member 14/Herpesvirus entry mediator
TNFSF10 : TNF superfamily member 10
TP53 : Tumor Protein P53
TrxG : Trithorax group
TSS : Transcription starting site

U

USP11/22 : Ubiquitin specific peptidase 11/22
UTX : Ubiquitously transcribed tetratricopeptide repeat, X chromosome (KDM6A)

V

VAF : Variant allele frequency
VEFS : VRN2, EMF2, FIS2, and Su(z)12

W

WHO : World health organization

X

XBP1 : X-Box binding protein 1

Y

YAF2 : YY1 associated factor 2

INTRODUCTION

Preface

Massive parallel sequencing of large cohorts of tumor samples has led to the identification of, not only recurrent and driver alterations, but also - and mostly - more passenger events, specific to certain tumor types. Deciphering genetic landscapes of cancers has led to a better understanding of the mechanisms used by cancer cells to drive their fate towards a malignant phenotype. Genetic changes, by definition *irreversible* and heritable through cell division, represent one of the most consistent molecular bases of tumorigenesis, ubiquitous through cancer diversity. Protein coding mutations, as well as transcripts arising from chromosomal alterations, thus represent surrogate targets in order to develop efficient drugs that are more adapted and tailored to kill or impair cancer cells than classical chemotherapies or radiation therapy. Several innovative therapies directed towards somatically mutated genes, identified through clinical Next Generation Sequencing (NGS), are approved each year, either as a standard of care or as a drug evaluated within the framework of clinical trials. In that regard, genetic events are sometimes given equal - if not more - weight than histological diagnosis. Data generated by NGS, and interpreted by physicians, to refine diagnosis or administer specific treatment to cancer patients, is generally referred to as *precision medicine*. As an ongoing concept, precision medicine is continually evolving, as understanding of cancer biology grows.

In spite of major advances achieved with the rapid rise of NGS, the understanding of cancer initiation or progression processes, as well as the mechanisms of sensitivity or resistance to therapies, remains incomplete. Genomic landscapes of myriad tumor samples noticeably showed that a high fraction of mutations reported in cancer affects genes involved in processes shaping the epigenetic status of the genome. These pathways refer to the regulation of gene activity with no alteration of their sequence, through multiple mechanisms including chromatin organization and composition. With this observation, a door was opened, triggering the investigation of yet unsolved questions when considered from a “classical” genomic point of view: a key for cancer transformation or progression is the modulation (i.e. activation or repression) of the expression of certain genes (i.e. oncogenes or tumor suppressive genes, either at the scale of a single gene or groups of genes) independently from the level of integrity of their coding sequence. Beyond the role of so-called epigenetics players in tumor physiopathology, it is now clear that they also represent important levers for the development of therapies.

A large number of genes, responsible for modifying DNA methylation or chromatin structure, composition or folding, have been identified as connected to cancer, whether mutated or not. Chromatin modifiers Polycomb group proteins (PcG) epitomize this category. In particular, PcG member Polycomb repressive complex 2 (PRC2) is a highly conserved chromatin-modifying enzymatic complex that methylates histone H3 on lysine 27 (H3K27me3), and plays a critical role in the maintenance of repressive transcriptional states *via* the deposition of H3K27me2-3 at the regulatory sequences of silenced genes. As such, PRC2 is a protector of cell identity, with many target genes involved in stemness, development and proliferation. Thus, it is not surprising that PRC2 elicits defects in the progression of cancer from diverse histological origins.

PRC2 alterations encountered in cancer and their consequences on H3K27me3 and on the chromatin landscape are multiple, yet unsettling to date as PRC2 can act alternatively as a tumor suppressor or, given the context, as an oncogene. Nonetheless, discoveries about PRC2 have been made at a rapid pace in this past decade, culminating in a recent milestone with the approval of a therapy targeting an activating mutation of PRC2 in patients with follicular lymphoma. However, some questions remain unanswered. This example illustrates the relevance of conducting comprehensive studies into this topic. The goal of this thesis is thus to provide a deeper insight into the role of PRC2 in cancer generally, and in follicular lymphoma in particular.

PART I PRC2 and its alterations in cancer

This first chapter will outline PRC2 structure and its main functions, while also providing an overview of how it can be altered in cancer, more specifically within the framework of alterations leading to its loss-of-function. Ultimately, some examples will illustrate how such alterations can be leveraged for therapeutic strategies.

I.1 Chromatin, Polycomb proteins and transcriptional repression

I.1.1 Background

Human genome is organized into a tridimensional scaffolding within the nucleus, chromatin: its 3.2 billion nucleotides are assembled together into a 2-meter-long DNA molecule that is wrapped around octamer of core histones to form nucleosomes, the structural units of chromatin, which are in turn stacked into the nucleus. This organization serves multiple purposes: from a physical perspective, genome is stabilized and protected from mechanical stress; and from a functional perspective, chromatin conformation determines accessibility to transcriptional machinery, thus modulating gene expression. Indeed, the mechanisms underlying such conformation aim to maintain, more or less faithfully according to the context, a specific pattern of gene expression through cell divisions. Influence of chromatin state is rather critical in a multitude of aspects of an organism's life, including embryonic development, establishment and maintenance of cellular identity, as well as programmed cell death.

Chromatin conformation varies with its composition, with DNA methylation and histone content, their post-translational modifications (PTMs) being important drivers to ensure its modulation and/or stability. These parameters contribute to a high and complex level or organization of genomic information, with no concomitant modification of the sequence of DNA itself and, therefore, fall into the field of epigenomics within its most widely accepted definition. The collection of these chromatin components and their modifications as a whole in a given state of differentiation of the cell is called *epigenome*. Cancer cells differ from normal cells in many ways as they progressively evolve and eventually ignore signals from the environment, hide from the immune system and spread away from their original location. Modification of the epigenome and specifically mis regulation of chromatin is one of the changes that can drive such an evolution.

Nucleosomes are made of a core histone octamer of approximatively 11 nm diameter around which a 147 base-pair (bp) long DNA section (representing a little less than 2 turns of the octamer) is wrapped. Each octamer consists of two repeats of four histones: H2A, H2B, H3 and H4. Histone H1 does not, strictly speaking, form the octamer, but instead corresponds to a histone linker that associates the nucleosome itself and both the entry and exit sites of DNA. The C-terminal part of the core histones protrudes outwards from the nucleosome, forming tails, which are subject to PTMs. These PTMs include, but are not limited to, acetylation, methylation and ubiquitination, which could impact chromatin environment regarding both the modified residue and the type of modifications. Indeed, nucleosomes are “spread” all along the DNA molecule every 200 bp forming a “beads-on-a-string” structure that is plastic, rather foldable, and further coiled into a helical 30 nm fiber, providing another layer of compaction. Chromatin thus consists in a succession of multiple layers of DNA compaction that reaches its highest level of condensation at the time of mitosis and meiosis (***Figure 1***) [1,2].

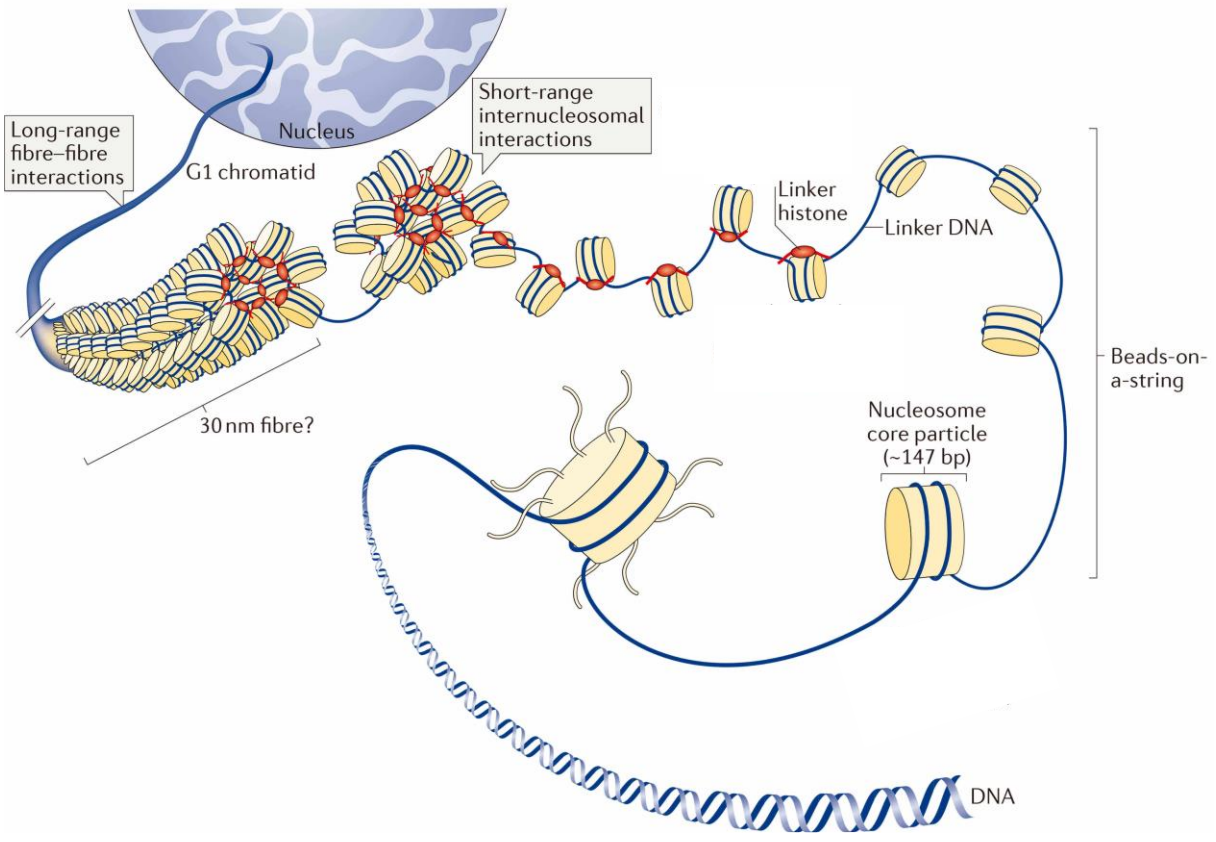


Fig 1: Multiple levels of chromatin folding, *adapted from Fyodorov et al, 2018* [3]

Such architecture is not clotted but highly dynamic and compartmentalized: in a nutshell, euchromatin is lightly packed and enriched with genes under active transcription, whereas heterochromatin is tightly packed, made of mostly silent regions. Heterochromatin can be divided into constitutive heterochromatin mostly composed of repetitive fractions of the genome (e.g. pericentromeres, centromeres, telomeres, etc), on the one hand, that is made stably inaccessible to transcriptional machinery; and, on the other hand, facultative heterochromatin, which can alternate from a condensed to an open state in order to allow transcription under certain circumstances [4]. This bimodal conception of chromatin organization is further refined into a more complex sub-structure with the constant amelioration of technologies specific to genome studies. For example, topologically associated domains (TADs) are self-interacting genomic regions, which means that DNA regions, localized within the same TAD, physically interact with each other more frequently than with sequences outside of the TAD [5]. Cohesin protein complex delimits chromatin loops and facilitates chromatin folding within the TADs. TADs are thought to participate to the genome physiology, in part through gene expression, with a dynamics of co-regulation of genes localized within the same TAD [6]. Chromatin tridimensional structure also contributes to gene expression regulation by favoring or restricting long-range interactions between common regulatory elements for a given genomic region. At a higher scale, nuclear organization can also modulate arrangement of various part of the genome. Every layer of chromatin ultrastructure, from the nucleosome to the entire chromosome territory, has been shown to be associated with cancer when altered.

I.1.2 Control of transcription by Polycomb machinery

Independently from the tri-dimensional nature of genome organization, transcription as a plain process follows a linear pattern where the template DNA sequence used for synthesis into RNA (coding and non-coding) opens up, allowing access to RNA polymerase II (POLII) at specific promoter regions (*initiation step*). POLII further synthesizes an RNA strand complementary to the template DNA strand in the 5' to 3' direction, while reading the template strand in the 3' to 5' direction (*elongation step*). The synthesized RNA chain briefly remains bound to the template strand, and is then removed from POLII upon reading of sequences, called t*erminators*, allowing the DNA to close back up and re-form the double helix (*termination step*) [7].

Transcription is a process that is under tight control and has to sharply respond to a multitude of internal and external stimuli. Both dynamics of chromatin compaction and interactions

between the genome and proteins, called *transcription factors* (TFs), participate in tailoring regulation of transcription.

TFs bind to DNA at specific target sequences with up to 10^6-fold higher affinity than for the remaining DNA, in order to activate (or, more rarely, inhibit) transcription. Activation of transcription by TFs is dependent on structural motifs on the DNA sequence, mostly localized at promoters or enhancers, in turn enabling accessibility and recruitment of POL II machinery *via* the assembling of the pre-initiation complex. Recruitment of POLII by TFs may be a direct mechanism or the last step in a cascade of several intermediate TFs activation. Activation of sequence-specific TFs are thus one of the most crucial and preponderant mechanisms of gene regulation in eukaryotic cells [8,9].

However, accessibility of the target sequences for TFs is not always optimal and transcription happened to be hampered by chromatin compaction. Indeed, the 3D state of chromatin at a given stage of differentiation or under the exposure of a certain stress can either promote or limit the access to transcriptional machinery. Chromatin conformation can not only determine the activated or repressed nature of a gene in time and space, but also plays a critical role in reversing transcriptional states as well as maintaining them over time and cell division, thus shaping the concept of transcriptional memory. To this end, a myriad of biological signals act together, either in cis- (e.g. DNA methylation) or in trans- (e.g. non-coding RNAs), in order to achieve such control on gene expression along with chromatin folding/unfolding and orchestrate genomic information inheritance. All of these mechanisms have been demonstrated to take part in one or many of the steps that lead a normal cell to undergo tumoral transformation while mis regulated.

One of the aims of this thesis is to provide a deeper understanding of how a chromatin landscape can be central to the development of certain cancer types – and more specifically, through the alteration of Polycomb Repressive Complex 2 (PRC2).

I.1.2.1 Polycomb Repressive Complexes

PRC2, along with PRC1, belongs to the Polycomb group proteins (PcG) that have been originally described in *Drosophila melanogaster* as well as its antagonistic complex Trithorax (TrxG): Both PcG and TrxG were shown to critically impact the antero-posterior axe during fly development, that is strictly under the control of *Hox* genes (coding for TFs) expression. A mutation leading to a change of phenotype characterized by ectopic sex combs in male flies,

was reported and named *esc.* Subsequently, a multitude of genes were identified as their inactivation altered the regulation of spatiotemporal expression pattern of homeotic genes *Hox* in a synergistic way, resulting in posterior transformation of the larvae along with somite segmentation. These genes were then named Polycomb [10]. Since then, orthologs have been identified in various species as PcG. PcG are highly conserved throughout evolution, in both a structural and functional way [11]. Today, PcG proteins are commonly defined in mammals by their constitutive interaction with one of the core members of either PRC1 or PRC2. Apart from rare exceptions, PcG play pivotal roles in development, pluripotency or differentiation as they participate in establishing and maintaining transcriptional memory [12].

Both PRC1 and PRC2 are multimeric enzymatic complexes and share as a main function to repress gene expression *via* PTMs of histone. PRC1 harbors a E3 ubiquitin ligase that catalyzes monoubiquitylation of histone H2A on its lysine residue 119 (symbolized H2AK119Ub, or H2AUb) [13]. PRC2 has a methyltransferase activity and is able to perform methylation (mono, di and tri) of histone H3 on its lysine residue 27 (H3K27me1/2/3) [14].

The paragraphs below will summarize the main structural characteristics (***Figure 2***) and known roles of both PRC1 and PRC2. Biology of PcG remains a field of investigation in constant evolution as previous chromatin studies are, in general, challenged and contradicted by more recent discoveries. One should therefore keep in mind that current knowledge about PcG might eventually be dismantled in the future.

I.1.2.1.1 PRC1

Various “blends” of PRC1 are described, based on which subunits assemble with each other. PRC1 complexes are divided into canonical complexes (i.e. composition similar to the originally described PRC1 complex) and variant complexes (i.e. with alternative, unusual functions, formerly named *non-canonical*). All PRC1 complexes are defined by the presence of either RING1A or RING1B E3 ubiquitin ligase.

Canonical PRC1 (cPRC1)

In addition to RING1A/B, one CBX (2/4/6/7 or 8) and one PHC (1/2 or 3) subunits, as well as SCM (1 or 2) and one of the two common factors in canonical and variant PRC1, PCGF2/MEL18 or PCGF4/BMI1, assemble together to form the canonical PRC1 (cPRC1).

cPRC1 relies on the property of CBX to bind to H3K27me3 *via* its chromodomain. The combination of subunits within PRC1 impacts both its recruitment to DNA and enzymatic activity. Hence, PRC1 E3-ligase activity is largely enhanced with the adjunction of PRC1 cofactors *in vitro,* compared with RING1A/B alone [15]. PHC2 has been shown to induce nucleosome oligomerization, and thereby restricts access to TFs [16]. Subsequently, PHC2, as well as the CBX subunit, play a central part in chromatin compaction, and PRC1's role in repression might only be partly dependent on its ubiquitination activity [17]. Indeed, not all of the mammalian CBX proteins contained in canonical PRC1 have proven their selectively bind to H3K27me3 [18]. Therefore, CBX proteins exert a spectrum of functions that remains to be elucidated.

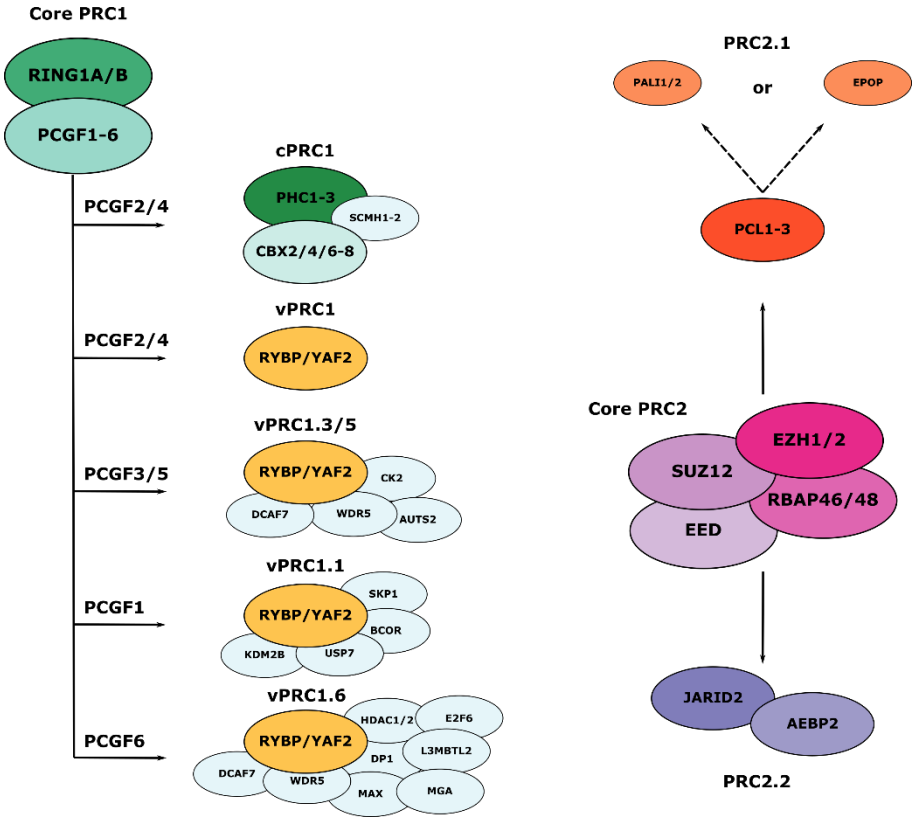


Fig 2: Polycomb Repressive Complexes 1 & 2, *representation of PRC1 inspired from Schuettengruber et al (2017)* [19]

Variants of PRC1 (vPRC1)

Variant PRC1 assembles in addition to RING1A or RING1B with RYBP or YAF2. Ultimately, vPRC1 encompasses one of the 6 PCGF proteins. PCGF2/MEL18 and PCGF4/BMI1 can be found in both cPRC1 and vPRC1, as mentioned above. Both RYBP and YAF2 proteins have an amino-terminal zinc finger domain that connects with H2AUb, and a carboxy-terminal RING1B-interacting domain [20]. In addition to the aforementioned factors, vPRC1 has been reported to be associated with a variety of chromatin factors.

Unlike their canonical counterpart, vPRC1 do not include CBX and do not require the recognition of H3K27me3 to dock to chromatin. Its ability to bind to chromatin is therefore independent of PRC2 activity (cf. paragraph I.1.2.1.2). Still, both RYBP and YAF2 have been shown to modulate enzymatic activity of RING1B *in vitro* with direct effect on the H2AUb level [21]. The respective role of each variant of PRC1 towards each other is not clear, but different forms of PRC1, canonical or variant, can occupy various loci, suggesting a cooperation of several PRC1 assemblies for transcriptional repression in a given cellular context.

Variant PRC1.1 modulates repression by protecting CpG islands from abusive DNA methylation

PRC1.1 is characterized by the presence of KDM2B/FBXL10 in which CxxC-zinc finger 1 domain is required for RING1B binding to unmethylated CpG islands in mESCs, as well as in the maintenance of this unmethylated state [22–24]. Limited DNA methylation at these loci favors a role for PRC1.1 in preventing them from forming a too stringently repressive state because of excessive DNA methylation.

The presence of PCGF1 in PRC1.1 is not required for the binding of RING1B to chromatin, but happens to be crucial for H2AK119 ubiquitination [25]. PRC1.1 (and specifically the PCGF1 subunit) is the PRC1 variant that is the most dependent on the stability of cPRC1 for binding to chromatin, as deletion of both *Ring1b* and *Ring1a* induces a significant removal of PCGF1 from chromatin in mESCs [26].

Variants PRC1.3 & PRC1.5 favor deposition of H2AUb and share a role with PRC2 in X chromosome inactivation

PRC1.3 and PRC1.5 structural similarities echo a certain level of functional complementarity [27]: In mESCs, deletion of either PRC1.3 or PRC1.5 does not substantially impact the amount of H2AUb. However, co-deletion of both complexes results in a noticeable decrease in H2AUb [25]. Forced recruitment assays performed in both mESC and in kidney cancer cells 293T led to conflicting conclusions about whether PRC1.3 ad PRC1.5 are associated with the deposition of either active or repressive marks [15,28].

PCGF3 and PCGF5 have also been reported to interact with *Xist*, a long non-coding RNA that plays a fundamental role in X chromosome inactivation [29]. PCGF3/5 double depletion is lethal with a more severe phenotype in female embryos. This implication of PRC1.3 and PRC1.5 in H2AUb-mediated X inactivation seems to be a unique feature of PRC1 complex biology, a phenomenon wherein PRC2 also plays its part.

Variant PRC1.6 is involved in germline specific transcription

Forced recruitment to chromatin showed that PCGF6 is required for binding to RING1B, and thus transcriptional repression, due to the contribution of L3MBTL2 and MAX-MGA, but independently from RYBP [26,30]. Of note, L3MBTL2 has been shown to favor chromatin compaction independently of histone PTM [31]. MAX-MGA is a putative TF that helps in PRC1.6 binding to the E-box motif. From an ontogeny point of view, PRC1.6 is noticeable as it binds to a set of promoters controlling genes implicated in spermatogenesis [32].

In sum, PRC1 assembling and orchestration are quite complex, and PRC1 mediated repression is the consequence of a tightly interlaced participation of the different members of the complex in deposition, propagation, or maintenance of different pools of H2AUb in a context specific manner.

I.1.2.1.2 PRC2

The core complex of PRC2, necessary for its catalytic function, consists in the association of EZH1/2, the Zinc-finger protein SUZ12, and the WD-40 protein EED, along with RBAP46/48, which also assembles to non-PRC2 complexes, such as LINC, NURF, NURD and SIN3 [14,33–35]. A cryo-electron microscopy-based representation of PRC2 is depicted in ***Figure 3*** [36]. PRC2 has fewer cofactors known to date than PRC1 and is proposed to come in fewer flavors. Nonetheless, two subtypes of PRC2, PRC2.1 and PRC2.2 are described to have distinct but partially redundant function.

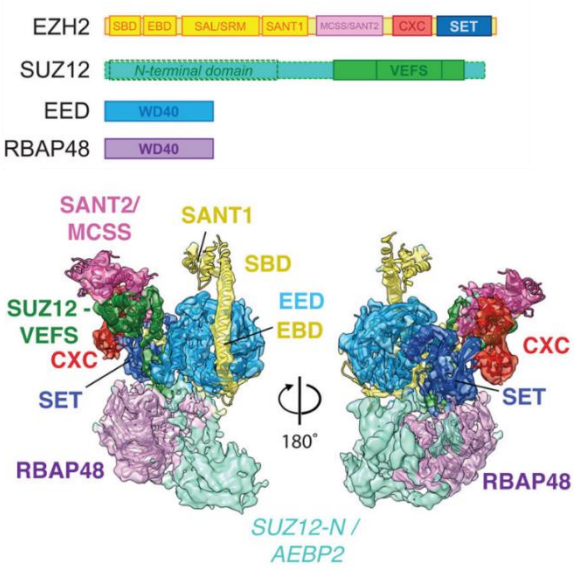


Fig 3: Cryo-EM structure of PRC2

Bar diagram of the PRC2 subunits and domains used for this representation (upper). Cryo-EM reconstruction of PRC2 at 4.6 Å with fitted crystal structures (lower).

Adapted from Poepsel et al (2018)

Core members of PRC2

Ortholog proteins EZH1 and EZH2 are characterized by a highly conserved catalytic SET [Su(var)3-9, Enhancer-of-Zeste and Trithorax] domain that is responsible for enzymatic activity of PRC2 - the main substrate being H3K27. EZH1/2, is only active with a fully assembled core complex. Even though EZH1 are EZH2 are functionally redundant, deletion of *Ezh2*, but not *Ezh1*, is responsible for a decreased in H3K27me3 in mESCs, underlying the higher efficiency of EZH2 over EZH1 [37,38]. Our lab showed that, given that the expression of EZH2 is correlated with the proliferative rate [39], EZH1 compensates for low EZH2 activity but only below a certain pace of cell proliferation [40]: in a highly proliferative rate, EZH2 is more expressed than EZH1. Several studies have reported PRC2 independent function for EZH2. Yet, to date, those

studies are scattered and no consensus has emerged regarding this so-called non-canonical function. I have, therefore, opted not to develop this particular aspect [41].

Another subunit, EED, is responsible for binding to H3K27me3 by means of its aromatic cage formed by WD-40 repeats. EED plays a critical role in allosteric activation of the complex. Indeed, loss of EED leads to a complete dismantle of PRC2 and, in turn, abrogates H3K27 methylation. EED also participates in the propagation of PRC2, as well as in the self-reinforcement of its activity [42,43].

In a similar fashion, SUZ12 is mandatory for the stability of PRC2 as its C terminal VEFS domain interacts with EZH2 and EED. However, unlike EED, complex-free SUZ12 is still able to bind to chromatin at CpG island rich loci, suggesting that its binding is not mediated by H3K27me3, and that it might have a PRC2 independent duty; but this remains unproven so far [44].

Mutual exclusivity of PRC2 variants

PRC2 variants are broadly classified as follows: PRC2.1 contains one of the three Polycomb-like paralogs (PCL1/PHF1, PCL2/MTF2 or PCL3/PHF19), and further consists of two mutually exclusive PRC2.1 variants: the EPOP-containing PRC2 and the PALI1/2-containing PRC2. PRC2.2 contains the defining subunits JARID2 and AEBP2 and, therefore, is often referred to as JARID2–AEBP2-containing PRC2. As puzzling as PRC1 can be, PRC2 is no exception to the rule and several studies have demonstrated that JARID2, initially described as an exclusive PRC2.2 cofactor, can clearly interact with at least one PCL (PCL2), questioning the dogma that distinguishes PRC2.1 from PRC2.2 [35,45–48].

PRC2.1

PCL proteins

Pcl triple KO (Pcl1/2/3) in mESCs shows a strong reduction of H3K27me3 deposition at PRC2 target genes, with no global depletion at the entire genome level. Each PCL protein plays its individual part in enhancement of PRC2 activity, yet no univocal mechanism of action is defined for each of these PRC2.1 members [49].

PCL1/2/3 proteins share a Tudor domain that binds to H3K36me3 *in vitro,* which might seem, at first glance, rather counterintuitive given the mutual exclusivity existing between H3K27me3 and H3K36me3 [50–53].

It has been proposed that the H3K36me3 demethylase NO66 and the PRC1.1 subunit KDM2B are recruited in a contemporary way with PCL3/PHF19, based on studies in mESCs. These demethylases eventually participate in the removal of H3K36me3 upon PRC2 recruitment, thus explaining why H3K36me3 is absent at PCL3/PHF19 and PRC2 binding sites in a static state: hence, the concept of a transient contact between PCL3/PHF19 and H3K36me3 would reconcile *in vitro* and *in vivo* observations, with a role for PCL3/PHF19 to fuse PRC2 binding [50,51].

PCL2, plays a major role in the maintenance of EPOP and PALI1 levels [49]. Its depletion in mESCs might lead to an impairment in differentiation *via* the indirect stabilization of pluripotent factors [54].

Moreover, PCL1/2 proteins would strongly promote the binding of PRC2 core subunits at unmethylated CpG-rich motifs in mESCs [55]. But this observation may not be true for all the CpG loci [49].

EPOP-containing PRC2 (a PRC2.1 variant)

The N terminus of EPOP contains a BC box, which interacts with both PRC2 and the elongin BC heterodimer and further fosters elongation. H3K27me3, as well as H3S28 phosphorylation, might inhibit EPOP and contribute to PRC2 release from chromatin. Moreover, EPOP plays a role in the maintenance of PRC2 targets with low expression [46,56].

PALI1/2-containing PRC2 (another PRC2.1 variant)

PALI1/2 share a globular PIP (PALI interaction with PRC2) domain in their C term regions and do promote PRC2 activity. PALI1 also contains another globular domain that interacts with G9A, a methyltransferase that deposits H3K9me2, as well as with deubiquitinases USP11 and USP22. Interestingly, promoters targeted by PALI1 correspond to loci that gain H3K27me3 in *AEBP2* null mESCs, suggesting that PRC2.2 might limit PALI-containing PRC2 activity [57].

PRC2.2, AEBP2-JARID2 containing PRC2

JARID2 and AEBP2 act synergistically and promote the catalytic activity of PRC2: their participation individually cannot achieve this. AEBP2 is important for JARID2 binding to the core PRC2, whereas JARID2 is required for maintaining normal AEBP2 levels in mESCs. JARID2 is able to interact with PCL2, while AEBP2 is not [45,49,58].

Initially, JARID2 was classified as a histone methyltransferase because of its JmjiC domain that is conferring, in general, such enzymatic activity. Yet no such capacities have been reported for JARID2. JARID2 is important for PRC2 core complex binding to chromatin, while enhancing its catalytic activity *via* H2AUb direct recognition, demonstrating the necessary interlace between PRC1 and PRC2 [58–61]. Moreover, JARID2 has been shown to be implicated in the initial *Xist*-induced X chromosome inactivation triggered by PRC2 [62].

AEBP2 is a finger zinc protein that is believed to have the capacity to foster PRC2 catalytic activity through promotion of PRC2 binding to the nucleosome. The short isoform of AEBP2, mostly present in embryonic tissues, might be more effective than the long one (in adult tissues) to promote gene silencing. Beyond its role in allosteric activation of PRC2 in duet with JARID2, AEBP2 binding to the β-sheet-rich domain of SUZ12 would actually help to overcome the H3K4me3 inhibitory effect on PRC2 activity [63,64].

PcG, other than PRC1 and PRC2, have been described such as Pho-RC, which is active in Drosophila [65–67], and PR-DUB. But the very definition of the latter as PcG does not appear appropriate in mammals [68]. Finally, EZHIP-containing PRC2 variant is a recent discovery made by several groups including our lab. EZHIP (a.k.a. cxorf67 or CATACOMB) inhibits methyltransferase activity and its expression is restricted to a subset of tissues (gonads, placenta, and certain parts of the central nervous system). In mice, deletion of EZHIP increases H3K27me3 and compromises female fertility [69]. Recent discoveries of EZHIP-mimicking mutation H3.3K27M, in a subset of aggressive pediatric brain tumors, are of major interest, and will be further discussed in the next paragraph of this first part of introduction (cf. I.2.2.2).

I.1.2.2 PRC1 and PRC2: recruitment, propagation and release of the marks

PRC1 and 2 preferentially bind to CpG rich regions, mostly unmethylated, close to the transcription starting sites (TSS) [56,70–72]. PRC1 and PRC2 binding patterns have been reported

to overlap, suggesting their complementary roles in silencing the same targets [73]. PRC1, at least partially, binds to DNA *via* its cofactors, unlike PRC2 that is instead recruited directly at unmethylated CpG.

Literature about the crosstalk between the mutual recruitment of one complex by another is abundant and contradictory in the hierarchy rules. Yet, the dogma that PRC1 recognizes H3K27me3 deposited by PRC2, thereby promoting its proper binding, has not always been verified. Indeed, PRC2.2 has been shown to recognize H2AUb prior to its activation [74,75]. I will not deeply discuss this point here. We might thus consider that both H3K27me3 and H2AUb stabilize PRC1 and 2 recruitment respectively, regardless of the intrinsic temporality of these events (***Figure 4***). Of note, lncRNA *Xist* might be involved in PRC1 recruitment, as mentioned previously [29].

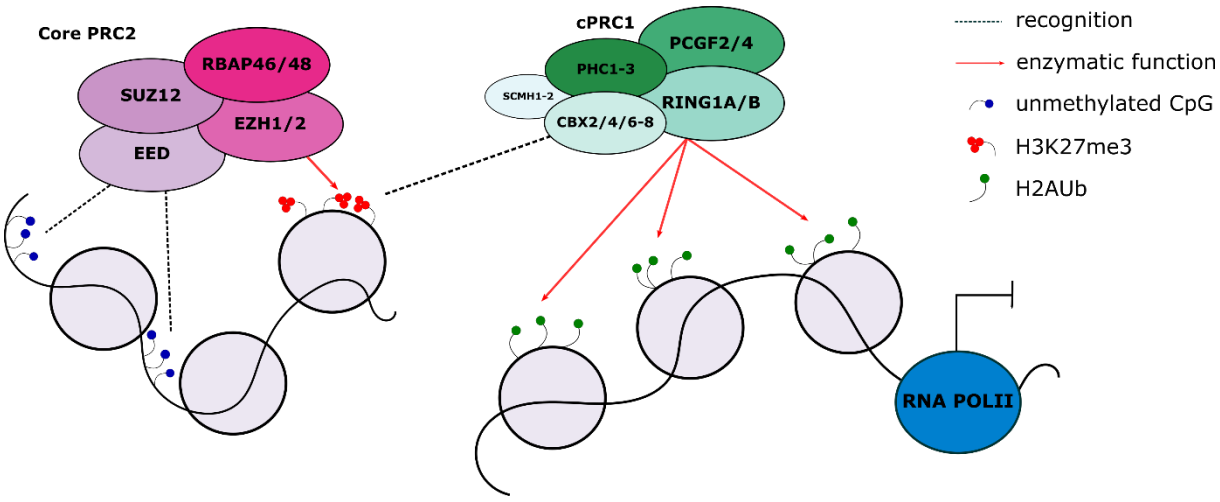


Fig 4: Recruitment and coordinate silencing activity of PRC1 and PRC2

Both PRC1 and PRC2 have self-reinforcement and self-propagation properties. PRC1 recognizes H2AUb through RYBP, and allosteric activation of PRC2 is mediated by the recognition of H3K27me3 by EED, before subsequent propagation [20,76]. Noticeable loss of EZH2 automethylation and impaired deposition of H3K27me3 by PRC2 beyond its nucleation sites upon H3K27 change of conformation, in that case termed *oncohistone*, is one of the driver alterations found in cancers that abrogates normal activity of PRC2.

Active transcription itself undoubtedly plays a role in refraining Polycomb machinery. Indeed, active marks, such as H3K4me3 and H3K36 me3, cannot colocalize with H3K27me3 on the same histone tail [64,77]. Moreover, it has been put forward that pre-mRNA removes PRC2 from chromatin at active genes, through an allosteric inhibition of EZH2 [78–80]. DNA methylation and PRC1/2 binding act in an antagonistic manner, and increased methylation limits H3K27me3 deposition in mESCs [81]. Finally, aforementioned EZHIP protein is also involved in counteracting PRC2 in a very context-specific manner [69].

I.1.2.3 Normal enzymatic activity of PRC2

Seventy percent of the genome appears to be methylated on H3K27: H3K27me1 decorates around 5%–10% of histones in mESC, and is associated with the gene bodies of actively transcribed genes [82]. H3K27me2 is by far the most abundant state of H3K27me, as it is present at the level of around 50% of histones. It is associated with repressed genes and absent from active enhancers. To date, its exact functions are not well understood. In the review by Conway *et al*, the authors depicted H3K27me2 as a “repressive blanket” that would protect the genes from an inappropriate transcription or enhancer activity [83–85]. H3K27me3 is enriched at 5%–10% of histones and is clearly associated with repression [86]. Overall, 70% of the genome is methylated on H3K27 by PRC2, which might seem antithetical with the preferred operational recruitment of PRC2 at CpG rich sites. This latter feature may be mostly the prerogative of H3K27me3 over H3K27me1 and H3K27me2, and remains consistent with the fact that PRC2 colocalizes with H3K27me3 enriched regions, but not at H3K27me1 and me2 loci. This observation might relate to a more transient time of residence to chromatin needed for PRC2 to perform the first steps of the methylation (0 to –me1 and –me1 to –me2) reaction in comparison with the last one (-me2 to –me3) [87,88]. Moreover, the necessity of a more stable and longer interaction between PRC2 and its target loci, to achieve the deposition of H3K27me3, might be a comprehensive key for explaining the oncogenicity of the mutation of *EZH2* observed, among other chromatin modifiers, in follicular lymphoma, to be discussed in detail below (cf. paragraph II.3.2.1).

Structural bases for PRC2 enzymatic activity consist of the following sequence: EED, upon binding to H3K27me3, induces a conformational change of EZH1/2. Indeed, EZH1/2 is wrapped around EED through a loop that is topologically located in the direct vicinity of a “pre-SANT (Swi3, Ada2, N-Cor, and TFIIIB) domain on EZH1/2 (see also ***Figure 3***). That pre-SANT domain is affected by this change in conformation and, in turn, modifies its flanking SANT

domain and the SET domain nearby, the latter being responsible for the catalytic activity of PRC2 as it carries the SAM-binding pocket [42,89–91].

To date, precisely *how* the deposition of H3K27me3 by PRC2 regulates transcription remains unclear, but some studies suggest a role for abortive transcripts in the recruitment of PRC2. In this way, PRC2 would behave mostly like a "shutter", recruited upon repression, in order to maintain a silenced state of genes by setting thresholds for their activation, rather than as an active process responsible for initiating repression itself [92,93].

In sum, PRC2 plays a crucial role in the maintenance of a repressive state of genes, mostly involved in proliferation, development and cell fate transition. Loss of PRC2 thus weakens the stability of cell identity. Its quintessential function for orchestrating the balance between differentiation and proliferation renders PRC2 a highly powerful process to divert - for a cell to transform into cancer.

I.2 PRC2 and cancer

I.2.1 Background

There is no one mechanism by which PCRC2 is connected to cancer and so studies have inevitably taken diverse directions to explore these connections [83,94–96].

In 2001, both Visser *et al* and van Kemenade *et al* reported an increased expression of EZH2 in non-Hodgkin lymphoma in association with increased cell proliferation [97,98]. Later, similar observations were made in a multitude of cancer types in which EZH2 was overexpressed (or amplified), consistently in association with a poorer outcome and various pejorative prognostic factors. *EZH2* was subsequently identified as a target of the E2F-pRB tumor suppressive pathway that is frequently deregulated in cancer [39,99]. At the time, it was hypothesized that EZH2 was required for proliferation. However, given the co-dependent expression of EZH2 with the continual cell cycle, the tendency for EZH2 overexpression in highly proliferative - thus more aggressive – tumors, might reflect a consequence of proliferation *per se* or of a more dedifferentiated state. Thus, in order to determine whether this apparent overexpression may be functional, it would have been preferable to study H3K27me3 enrichment, rather than *EZH2*

transcript level. However, numerous studies reporting high levels of EZH2 in aggressive tumors arising from various tissue types, have been published since then [100]. The question of the overexpression of EZH2 under such circumstances will not be further developed here, given the scope of this thesis - to focus on structural alterations of PRC2.

In parallel with these reports, qualitative alterations of PRC2 subunits, mostly mutations, have been described in diverse types of tumors, leading to various, sometimes opposite consequences in terms of enzymatic activity of the complex, along with the transcriptomic effects that ensue (***Figure 5***). Indeed, one crucial aspect of altered PRC2 in cancer is that it can play either the role of a tumor suppressor, or of an oncogene in a context-dependent manner. Given the nature of the classical target genes of PRC2, some common traits of its implication in oncogenesis may still be emphasized, regardless of the context. Thus, tumor suppressors cyclin-dependent kinase inhibitors (CDKN), master regulators of cell cycle, have been reported to be mis regulated upon PRC2 alterations, be it through tumor suppressive or oncogenic mechanisms. Hence, repression of *CDKN1A* in *EZH2* mutant B cell lymphoma appears to be critical in germinal center lymphomagenesis. PRC2-mediated repression of CDKN2A is necessary for proliferation of rhabdoid tumors, and a combinatorial knock out of both *Eed* and *Cdkn2a* fosters cell growth in both leukemic MLL-AF9 cells and HSCs [101,102]. Noticeably, loss of *CDKN2A* has been reported to be a precursor event of Malignant peripheral nerve sheath tumors (MPNST), upstream from the loss of *EED* or *SUZ12* (cf. next paragraph) [103]. *CDKN* deregulation is one example of a common explanation for the role of PRC2 in cancer. However, this only applies to some tumors.

In general, we might consider that the deregulation of PRC2 results in a reduction in its function of stabilizer of cell identity, either favoring some stemness traits along with enhanced proliferation, or stochastically inducing both de repression and silencing of myriad genes, which in turn results in increased genomic instability, or both.

The following paragraphs will illustrate how cells more specifically hijack PRC2 activity according to the context. Indeed, loss-of-function mutations of PRC2 shape a spectrum that is highly heterogenous at both the clinical and the molecular levels. Increased activity of PRC2, observed upon *EZH2* mutation, often referred to as a gain-of-function, observed in B cell lymphoma, will be the focus of the second part of this introduction.

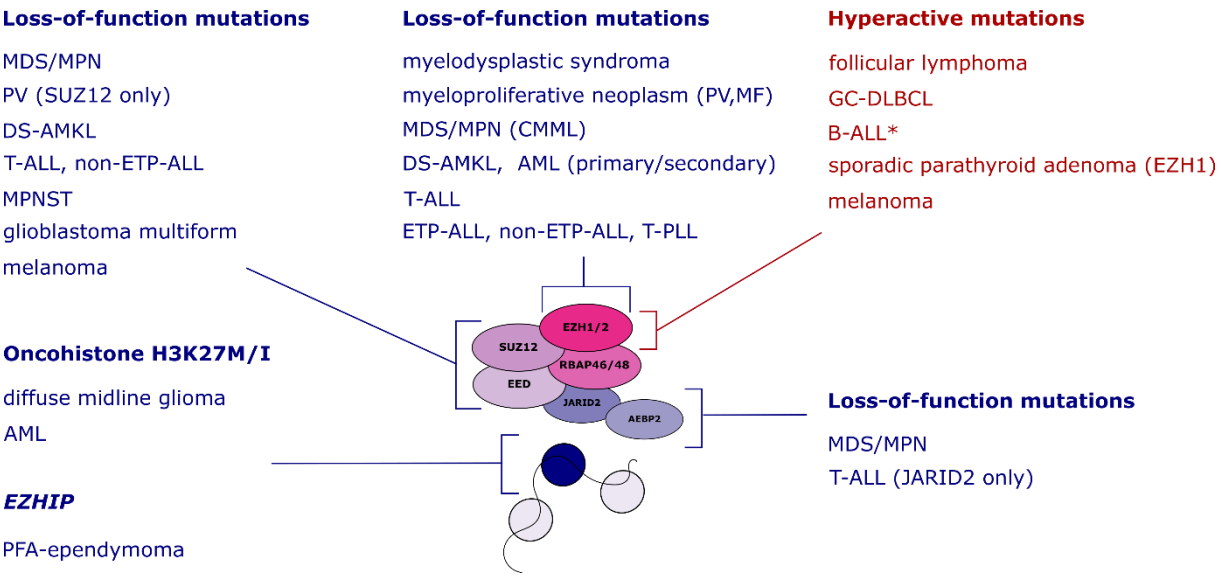


Fig 5: Types of alterations of PRC2 encountered in cancer

*MDS/MPN myelodysplastic syndrome/myeloproliferative neoplasm, PV polycythemia vera, DS-AMKL acute megakaryoblastic leukemia associated with Down syndrome, T-ALL T cell-acute lymphoblastic leukemia, ETP-ALL early T cell precursor ALL, MPNST malignant peripheral nerve sheath tumors, MF myelofibrosis CMML chronic myelomonocytic leukemia, AML acute myeloid leukemia, T-PLL T cell prolymphocytic leukemia, GC-DLBCL germinal center diffuse large B cell lymphoma, B-ALL B cell-ALL *patient-derived cell lines, PFA posterior fossa group A*

I.2.2 Loss-of-function mutations in genes encoding PRC2

I.2.2.1 Malignant peripheral nerve sheath tumors

Malignant peripheral nerve sheath tumors (MPNST), an aggressive subtype of soft-tissue sarcoma, arise in three different clinical settings: within the context of Neurofibromatosis-type I (NF-I, characterized by germline monoallelic deletion of *NF1,* and associated with multiple neurofibroma (the benign counterpart/precursor of MPNST), sporadically or as a secondary effect of radiotherapy. Regardless of the clinical background, PRC2 alterations, more specifically loss-of-function, have been reported to be highly recurrent in this cancer type, with either *SUZ12* or *EED* mutation found at rates ranging from 70% in NF1-associated MPNST to 92% of sporadic tumors [104]. More rarely, similar alterations can be observed in melanoma and glioblastoma multiforme. EED loss-of-function can relate to frameshift or splice site mutation along with loss of heterozygosity. As for *SUZ12*, both homozygous and heterozygous deletion have been reported. Interestingly, both altered subunits appear mutually exclusive and no other PRC2 member is ever altered in MPNST [40]. In this context, PRC2 unambiguously plays

the role of a tumor suppressor and logically, PRC2 deficient MPNST display a major decrease in H3K27me3 (***Figure 6***). This feature is consistent enough as a criteria that reduced expression of H3K27me3 is fully part of histological diagnosis of malignancies for peripheral nerve tumors since neurofibroma, in theory, do not harbor PRC2 alteration [105–107].

PRC2 impairment has a straight forward impact on gene expression, indeed, a signature of 480 genes robustly separated PRC2 WT from PRC2 deficient cases [104]. From an ontogenetic angle, the genes upregulated upon PRC2 loss-of-function are classical PRC2 targets (i.e. homeobox, genes development and morphogenesis genes). Moreover, H3K27me3 ChiP-qPCR analyses confirmed the direct link between PRC2-related silencing and abrogation of H3K27me3 deposition in this process. Furthermore, in an *Nf1*$^{+/-}$; *Trp53*$^{+/-}$ mouse model, heterozygous loss of *Suz12* significantly accelerated the development of MPNST, underpinning the synthetic lethality mode followed by NF-I associated MPNST *via* Ras mediated mechanism [108]. Of note, the close genomic proximity between *NF1* and *SUZ12* might explain why the rate of *SUZ12* lesions is higher than altered *EED* in *NF1*-associated MPNST. Restoration of proficient PRC2 in MPNST cells objected re-established deposition of H3K27me3 at target genes, along with decrease of cell growth. This asserts the unequivocal role of PRC2 loss-of-function in tumorigenic process of MNST. Potential therapeutic implications of these latter findings will be discussed in the following section.

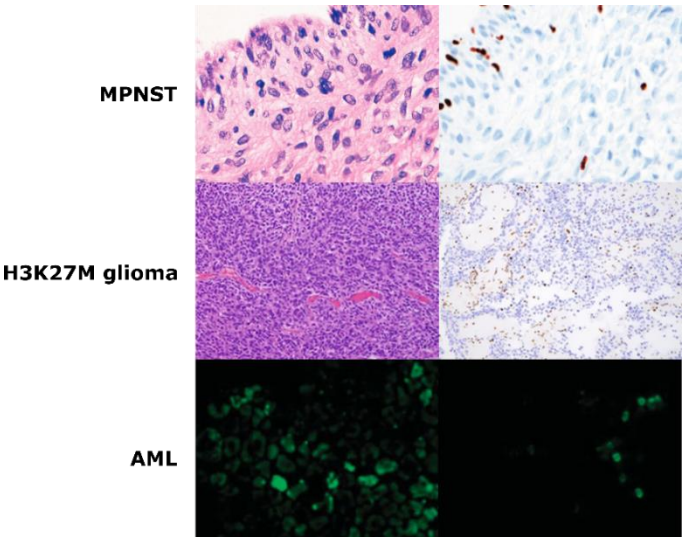


Fig 6: H3K27me3 reduction in cancers with PRC2 loss-of-function

Upper panel: H&E (left) and IHC H3K27me3 (right) in MPNST. Positive staining on lymphocytes and endothelial cells, adapted from Mito et al (2017); Middle H&E (left) and IHC H3K27me3 (right) in H3K27M glioma, adapted from Marchione et al (2019); Lower panel: H3K27me3 level assessed by immunofluorescence in AML cells from EZH2 WT (left panel) and mutant (right panel) cases, adapted from Khan et al (2013)

I.2.2.2 Oncohistone H3K27M and *EZHIP* in pediatric brain tumors

A spectrum of high-grade glioma diagnosed in children and young adults is characterized by alterations of PRC2 activity that relate to mechanisms different from a mutation of PRC2 subunits themselves: Recurrent somatic mutations in the substrate of PRC2 histone H3 - in this case, sometimes referred as *oncohistone* - are found at various residues (mostly lysines) of the histone tail, depending on neuroanatomical specificity. PRC2-related pathogenicity in pediatric brain tumors was recently extensively reviewed by Krug *et al* [109]. H3K27M variants (H3.3K27M, H3.1K27M) were reported in up to 85% of what were formerly called *diffuse intrinsic pontine glioma* (DIPG). DIPG were designated in the 2016 World Health Organization (WHO) central nervous system classification under the appellation *H3K27M diffuse midline glioma*, a highly lethal tumor type arising in thalamus, midbrain, and occasionally, the spinal cord. [110–112] Even though only the authentication of the mutation is a diagnostic criteria, immunohistochemistry studies consistently shows a reduction in H3K27me3 levels [113,114]. Interestingly, H3K27M oncohistone inhibits PRC2 in a negative dominant fashion, as no more than 10% of the whole pool of proteins coding for H3, a blend of several isoforms, happens to be mutated [115–117]. H3K27M hampers both H3K27me2/3 deposition and EZH2 automethylation and impairs allosteric EED activation [118]. However, it has been shown that residual H3K27me3/me2 deposition on chromatin of H3K27M-expressing cells remains. The mechanism by which PRC2 activity is partially abrogated is indeed not yet fully unraveled [119]. Some have authors suggested a sequestration of PRC2 by oncohistone at the nucleosome, resulting in an increased biding of PRC2 to chromatin [120]; and others, that the oncohistone does not modify affinity between H3 and PRC2 [121,122]. Finally, it has been proposed that it either retains PRC2 only at the sites with strong affinity [123] or excludes it from newly generated H3K27M-K27ac nucleosomes. Indeed, increased H3K27ac levels have been reported in H3.3K27M mutant cells [124]. Recently, it has been reported however that H3K27ac is focally lost, along with a repression of neurodevelopmental genes upon increased affinity between PRC2 and H3.3K27M at these loci [125]. Harutyunyan *et al* showed that oncohistone does not impair PRC2 docking at its nucleation sites but instead, limits the self-propagation of H3K27me3, which - at the transcriptomic level - translates into the de-repression of genes with low expression and repeats elements [119].

Thus, in addition to a clear loss-of-function of PRC2 in the presence H3K27M, the hypothesis of a residual, still oncogenic, H3K27me3 deposition at specific CpG islands exists. The mark would play a crucial role in tumor progression and could thus be targeted by EZH2 or BET bromodomains inhibitors. This point will be addressed in the following paragraph, which focuses on therapeutic approaches for cancers with PRC2 loss-of-function [123,126].

In sum, as observed in all PRC2 altered malignancies, oncohistone H3K27M induces a redistribution of other histone marks (H3K27me1, H3K27ac, H3K36me3) alongside H3K27me3 and might also generate an imbalance between PRC2 and PRC1 activities, which must contribute in many ways to the oncogenicity induced by the mutation [124,127–129]. Of note, as in the aforementioned example of MPNST, co-occurring mutations are found to foster H3K37M-related tumorigenesis, such as p53 deletion or frequent activating mutation in growth factor receptors crucial for brain development such as *ACVR1* [111,130].

Beyond the entity of H3K27M diffuse midline glioma, H3K27M is found in around 5% of PFA-ependymoma, another aggressive brain tumor in children. Interestingly, the 95% other cases are characterized by the overexpression of EZHIP, a PRC2 antagonist of recent discovery normally expressed mostly in (but not restricted to) gonads and placenta (cf. paragraph I.1.2.1.2) [69,131]. In PFA-ependymoma, EZHIP mimics the mutant H3K27M tail and produces similar effects to the oncohistone itself on PRC2 recruitment and binding to its nucleation site, with no incorporation of EZHIP to chromatin [132,133]. Mechanisms underlying overexpression of EZHIP in the very context of PFA-ependymoma are unclear and possibly involve loss of its promoter 5mC and responses to hypoxia [131,134,135].

I.2.2.3 Hematological malignancies

The intrinsic nature of Polycomb group proteins to modulate the transcription of genes involved in commitment for cell identity and proliferation is exacerbated in hematopoietic stem cells (HSCs), where they play a crucial role in the maintenance of hematopoiesis. Thus, deregulation of one or several members of the PRCs seriously risks leading to malignant transformation of hematopoietic cells at various stages of their development, from early progenitors to fully mature and differentiated cells. Both myeloid and lymphoid can be affected. Moreover, depending on the lineage and the cell stage at which the tumoral phenotype arises, the mutated subunit and the consequences of the alteration vary widely. Activating point mutation of the SET domain of *EZH2* reported in FL will be described in detail in the next chapter. In contrast, loss-of-function mutations have been described in all the core members of PRC2 in blood cancers (as well as in PRC1 subunits) [136].

In MLL-AF9 mice, a mouse model recapitulating human acute myeloid leukemoid (AML), *Ezh2* alternatively plays a role as a tumor suppressor during the induction phase or as an oncogene in the maintenance phase, illustrating the multifaceted nature of *EZH2* in this context [137–139]. In myelodysplastic syndrome/myeloproliferative neoplasm (MDS/MPN), PRC2 is most frequently

targeted through loss-of-function mutations of *EZH2* [140]. *EED* and *SUZ12* mutations are rarer but still reported (review in [136]). The first hotspot mutated region of *EZH2* is the pre-SANT domain, but the SET domain can be affected (***Figure 3***). In MDS, loss of the 7q chromosome, where *EZH2* is hosted (7q36), is recurrent [141]. Moreover, *EZH2* inactivating mutations are frequently found in AML arising from transformed MDS and associated with poorer prognosis, especially when bi allelic. Such an event is less frequent in *de novo* AML, suggesting a late contribution of the mutant to the natural history of MDS rather than a driver alteration in this context. Indeed, *EZH2* belong to a mutational landscape, including *TET2*, *RUNX1* and *ASXL1,* that are all other important chromatin factors [140,142]. Loss of *Ezh2* has been shown to cooperate with hypomorph *Tet2,* or mutant Runx1, to promote MDS and MPN in mice [143,144]. This cooperation might echo with the fact that deletion of *Ezh2* in HSCs results in a skew toward myeloid lineage repopulation, a background upon which additional mutations allow cancer transformation [138]. One interesting finding, which highlights the tumor suppressive role of *EZH2* in MDS, is the upregulation of pro-inflammatory IL6 in RUNX1S291fs/*Ezh2* null HSCs in mice [144]. Moreover, production of inflammatory cytokines, such as S100a8/ S100a9, have been reported upon *Ezh2* loss [145]. This illustrates how impaired PRC2 might influence tumor microenvironment in MDS.

EZH2, *EED* and *SUZ12* loss-of-function mutations are present in a subset (up to 15%) of T cell acute lymphoblastic leukemia (T-ALL). The mechanism underlying oncogenesis here is again not fully understood but a competition between oncogenic NOTCH1 activation and loss of PRC2 has been proposed as playing a central role in T-ALL [146]. P53 deletion has also been reported to foster leukemic transformation along with *Ezh2* loss through a mis regulation of DNA methylation at key T cell developmental genes [147].

In aggregate, PRC2 loss-of-function alterations in blood malignancies are diverse and their consequences highly context-dependent, relying on the very nature of the hematopoietic tissue. It remains puzzling to notice that oncohistone H3K27M (and H3K27I) (cf. previous paragraph) are found in rare cases of AML, with a potential role in promoting tumorigenesis in RUNX1 deleted AML mice [148,149]. Thus, even though there is a strong reliance in cell identity for altered PRC2 in cancer, examples exist of breaches between selective advantage processes that took over in one cell type or another.

I.3 Targeting PRC2 loss-of-function in cancer

Based on the observation that PRC2 alterations are recurrent in a wide range of tumors and result in the deregulation of large sets of genes, targeting these alterations appeared an obvious path to follow. The tumor suppressive nature of PRC2 however, represents a challenging hurdle. Main strategies developed in regard to the specific context are discussed below.

I.3.1 Malignant peripheral nerve sheath tumors

As mentioned earlier, MPNST represent a prototypical example of full loss-of-function of PRC2 in cancer. In terms of therapeutic management, surgery is necessary, whenever possible, according to the current standard of care, but MPNST grows and disseminates fast. Anthracycline-based therapy is the front-line therapy of choice for patients with unresectable, locally advanced, or metastatic tumors, but response rates remain poor [150]. As shown in experimental models, restoration of PRC2 proficiency has a positive impact on cell growth limitation, but no such feasible strategy can be considered in the clinical setting. However, consequences of deficient PRC2 might represent actionable targets. Indeed, the chromatin landscape of MPNST is perturbed beyond abrogation of H3K27me3 and in formalin-fixed paraffin-embedded (FFPE) tumor samples with PRC2 loss-of-function, Wojcik *et al* showed, similarly to what is observed in cell lines, an enrichment of H3K27ac [151]. Newly acetylated loci can be targeted by the bromodomain and extra terminal domain (BET) family proteins (e.g., BRD protein family), which in turn recruit transcription factors and modulate the expression of oncogenic genes, such as Ras signaling genes or enhancers [108,152]. Similar observations were made with the report of upregulation of BRD4, also a member of the BET family, in MPNST mouse model [153]. Consequently, JQ1, a bromodomain inhibitor, induced the silencing of genes normally repressed by PRC2. Furthermore, JQ1, in synergy with the MEK inhibitor PD901, promotes cell death and triggers tumor regression in a mouse model in the pioneer study aforementioned by De Raedt et *al* [108]. Thus, a phase II clinical trial evaluating the BET inhibitor CPI-0610 in MPNST (NCT02986919) was initiated but had to be interrupted due to poor enrollment [154]. In line with a largely altered chromatin landscape in MPNST, HDAC inhibitor-induced autophagy has been considered *in vitro* and *in vivo*, prior to the discovery of the existence of a subset of PRC2 deficient MPNST [155].

An interesting finding form the translational approach developed by Wojcik *et al* was the repression of antigen presentation genes in PRC2 deficient MPNST, most likely as a

consequence of indirect mechanisms. This study thus suggests a role for PRC2-related changes in immune evasion. Knockdown of *NSD2*, coding for a methyltransferase responsible for H3K36me2, as well as DNMTi (deficient PRC2 MPNST had increased DNA methylation) and HDAC inhibitor (HDACi) restored MHC expression in cells and induced interferon pathway expression, similar to what is observed with PRC2 restoration [151].

These promising observations epitomize how a PRC2 loss-of-function might affect other actionable pathways (i.e. H3K27ac), and therefore pave the way for innovative therapeutic strategies, here targeting the immune vulnerability of PRC2 deficient MPNST, or enhancer-driven differentiation elsewhere [152].

I.3.2 Oncohistone H3K27M in pediatric brain tumors

Some of the strategies considered for the treatment of PRC2 deficient MPNST are also being exploited in H3K27M mutant pediatric gliomas as both pathologies share a profound refurbish of H3K27me3 deposition pattern, leading to potentially actionable targets.

For example, BET inhibitors have proven their efficiency in H3K27M glioma cell lines and PDX based on the similar observation that not only H3K27me3 deposition is reduced, but also H3K27ac is increased upon the oncohistone expression, rendering the cells sensitive to bromodomain inhibition [126]. Moreover, a global increase of H3K27ac deposition might have implications for antitumor immunity with the abnormal expression of endogenous retroviral elements [124]. Polyacetylation, at adjacent residues, blocks the interaction of PRC2 with H3K27M. This is overcome with the use of HDACi panobinostat in both cell lines and PDX models and, to date (August 2021), 8 phase I or phase I/II trials enroll high grade glioma patients for HDACi evaluation, either alone or in combination [156]. Other interesting strategies can be mentioned, such as the inhibition of CDK7-mediated abnormal transcription resulting from H3K27ac deposition [157]. As for AML and multiple myeloma (cf. next paragraph), the loss of H3K27me3 might be mitigated through use of the K27 demethylase JMJD3 inhibitor GSKJ4. Such a strategy was indeed associated with improved survival in H3K27M PDX, but not in WT tumors [158].

One intriguing point to discuss here is the proposition for the use of EZH2 inhibitors that stems from studies showing that tumor growth relies on partial, yet crucial, retention of PRC2 activity in the presence of H3K27M [123,125,126]. Retained H3K27me3 peaks at tumor suppressor loci would represent a vulnerability for the cell toward EZH2 inhibition. Indeed, Brien *et al* reported, in the context of H3K27M isogenic neural stem cells, that specific repression of

neurodevelopment genes can be overcome by EZH2 inhibitor. Elsewhere, H3K27M glioma cells did not respond to such therapies [159].

Finally, exciting strategies based on immune targeting of H3K27M have been investigated, with the development of a peptide vaccine [160], adoptive transfer of TCR-transduced T cells [161] and chimeric antigen receptor (CAR) T cells, targeting tumor-associated antigens and showed promising results in mice [162,163].

I.3.3 Hematological malignancies

In a similar way with the two previous examples, not targeting PRC2 itself, but the downstream consequences of its impairment described in hematological malignancies, may represent a rationale for developing innovative strategies. Moreover, in the context of blood cancers, PRC2 alterations are often reported to be associated with chemoresistance.

Göellner *et al* showed, for instance, that loss of *EZH2* causes acquired drug resistance to tyrosine kinase inhibitors (TKI) and cytotoxic drugs in AML *via* different mechanisms, among which *HOX* genes de repression [164]. In mice, they showed the positive impact of targeting HoxA9 to restore sensitivity to TKI and, in line with these observations, a number of HOXA9 inhibitors, such as HTL-001, are planned to be soon under the scope of clinical evaluation for patients with AML [165]. Moreover, decreased EZH2 expression in resistant AML cells was also associated with a CDK1-dependent phosphorylation of EZH2 at Y487. This interaction, stabilized by Heat Shock Protein 90 (HSP90), led to the proteasomal degradation of EZH2. HSP90, CDK1 and proteasome inhibitors prevented this degradation and, in turn, restored sensitivity to TKI. Finally, patients with reduced EZH2 levels responded to the adjunction of the proteasome inhibitor bortezomib to the standard therapy consisting in DNA synthesis inhibitor cytarabine. Subsequently, restoration of the EZH2 protein level might be a potent strategy to overcome drug resistance in AML [164].

As mentioned in the previous paragraph, inhibition of H3K27 demethylases might provide promising results upon PRC2 loss of activity. Li *et al* showed that the use of GSKJ4 resulted in increased H3K27me3 levels, reduced proliferation and disease progression in a humanized AML mouse model [166]. Ezponda *et al* reached the same conclusions in multiple myeloma cells, and unveiled an unexpected increase in sensitivity of *UTX*/*KDM6A*-mutant (another histone lysine demethylase) cells to EZH2 inhibition. The rationale for the use of such a counterintuitive combination remains to be clarified [167]. Conversely, EZH2 overexpressing CML leukemia stem cells (LSC) have been reported to be sensitive to EZH2 inhibition, along with the use of TKI,

which might be of major relevance given that LSC are known to drive tumor relapse or resistance to TKI alone [168].

Mutational loss of *EZH2* might also be targeted under the concept of synthetic lethality. For instance, Gu *et al* have identified a potential vulnerability upon such loss: BCAT1, that catalyzes the transamination for branched-chain amino acids, is abnormally upregulated because of *EZH2* impairment in mice models and human myeloid neoplasm samples, and might represent a potential candidate as, in mice, *Ezh2*-deficient cells are more sensitive to BCAT1 inhibition than *Ezh2* wild-type cells [169].

Finally, not only the myeloid, but also the lymphoid PRC2 deficient malignancies can be triggered at the level of their epigenome as DNA hypermethylation observed in *Ezh2* KO T-ALL cells are responsive to hypomethylating agent decitabine [147].

Overall, even though strongly related to cell identity, the effect of PRC2 alterations encountered in cancer do actually share a substantial reprogramming of chromatin landscape (H3K27ac, H3K27me2, DNA methylation, etc…). Downstream effects of this reprograming result in a spectrum of attractive targets to overcome impaired PRC2 in tumors.

PART II *EZH2* mutation in follicular lymphoma

The purpose of this second part of my introduction is to provide an overview of the current knowledge about follicular lymphoma with an emphasis on the role of point mutation of *EZH2* in this cancer type.

II.1 Background

Follicular lymphoma (FL) is a malignant, mature B cell proliferation/neoplasm diagnosed with an incidence of approximately 2 in 100 000 in Europe, thus representing the most frequent indolent lymphoma subtype (~70%) [170,171]. FL is, by definition, composed of small B cells harboring morphological, immunophenotypic and, to a certain extent, functional characteristics of germinal center (GC) normal B cells (i.e. centrocytes and centroblasts). Therefore, FL cells are considered to recapitulate the state of differentiation observed in the GC of normal lymphoid follicles. From an architectural point of view, FL cells form tri-dimensional nodular structures that resemble normal lymphoid follicles in most cases. Consequently, the term *follicular* has been used to designate this subtype of lymphoma. Arising mostly in adults, FL can also be observed, in rare cases, in children or young adults. Thus, in addition to some other variants of FL, updated 4th edition of the WHO classification of lymphoid neoplasms defined the so-called pediatric-type FL, that is characterized by a high frequency of grade 3 and is associated with an excellent prognosis [172].

This thesis will only focus on so-called classical adult FL. Thus, my introductory paragraphs, as well as the results and discussion paragraphs will refer solely to non-pediatric-type FL.

II.1.1 Epidemiology

With all subtypes included, lymphomas are the seventh most frequent cancer worldwide. FL represents the second most frequent lymphoma subtype (diffuse large B cell lymphoma – (DLBCL) being the most frequent) and the first indolent one. Around 2,500 new cases are diagnosed each year in France [171]. Median age of diagnosis is 63. Men are slightly more impacted than women (sex ratio 1.2:1). FL incidence increased from the 1970s up until the early 2000s, undergoing a gently declining trend in the past two decades (***Figure 7***) [173].

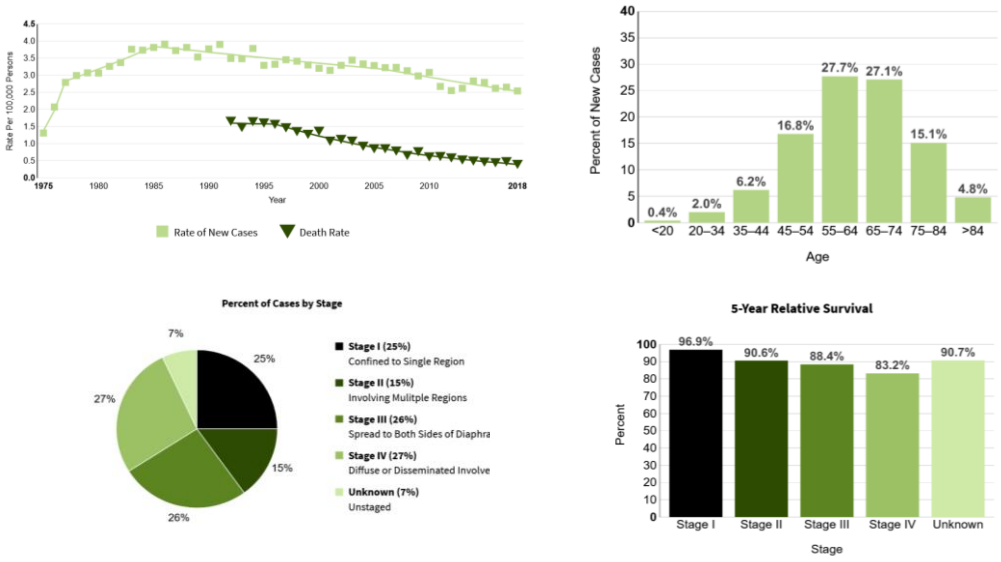


Fig 7: Follicular lymphoma: stat facts

New cases & death rate per year (upper left), percent of new cases per age group (upper right), percent of cases and 5 year survival per stage (lower), adapted from SEER database (USA, update august 2021) [173]

To date, no one cause has been put forward as responsible for the onset of FL. However, inherited susceptibility has been suggested by several observational studies, given the statistically increased association between a family history of non-Hodgkin lymphoma (FL or non-FL) and the individual risk of developing FL [174,175]. More refined studies further reported a noticeable association between FL and the presence of a SNP closely related to a psoriasis susceptibility region in chromosome 6 [176]. Most importantly, a pioneer GWAS study led by Wang *et al* showed, in a cohort of more than 2000 cases of FL, an increased risk of disease in association with the overall increase in the number of homozygous HLA class II loci [177].

Beyond genetic susceptibility, many studies have tended to correlate FL risk with exposure to various compounds or occupational and/or environmental factors, but the supporting evidence reported in these studies remains inconsistent [178–181]. A large metanalysis of 19 observational studies including 3530 cases of FL in comparison with more than 22000 non-FL cases in total first confirmed that family history of FL increases one's individual risk to develop FL [182]. Moreover, young adults with high body mass index as well as women with Sjogren syndrome have a higher risk of developing FL. Conversely, patients with an atopic history (asthma, hay fever, allergy), a history of blood transfusion, sun exposure, as well as bakers, milers and university teachers appear to be less at risk of developing FL. Overall, the authors confirmed the absence of a robust causal link between one determined environmental/occupational factor

and the risk of FL, and further emphasized the multifactorial etiology of FL. Additionally, no shared genetic risk was found in the GWAS study between FL and autoimmune disease by Din *et al.* [183]. In the 2000s, however, not only observational (still retrospective) but also translational studies, have highlighted the putative role of agricultural pesticides exposure in clonal expansion of B cells harboring t(14;18) and, in turn, increased risk of FL [184,185].

In sum, in spite of the absence of a straight forward explanation for the onset of FL, it seems that both population-based and individual genetic considerations, along with exposure factors, point toward the role of a challenged immunity as a modulator of the individual risk of developing FL.

II.1.2 Clinical aspects and course of follicular lymphoma

At the time of diagnosis, the general status of the patient is usually preserved and most of them typically present with barely symptomatic superficial lymph nodes (cervical, axillary, femoral, inguinal areas). Although uncommon, extra nodal sites may be affected at first onset (gastro-intestinal tract mostly). Clinical systemic features, typically referred to as *B symptoms* (i.e. fatigue, fever, night sweats, and weight loss), are often missing or, if present, may raise suspicion of histologic transformation (hereafter referred to as transformed FL, tFL). In cases of symptomatic disease, the first symptoms, other than general symptoms, may be related to the slow growth of lymph nodes in deep areas (e.g. retroperitoneum, mesenteric, or iliac areas). Primary mediastinal involvement is not usual for FL. The bone marrow is involved in up to 60% of these cases, explaining the relatively high rate of stage IV disease at diagnosis in the Ann Arbor classification, regardless of the indolent typical course of FL (***Figure 8***).

Ultimately, relapse is the general rule for almost all patients. Furthermore, overall 25% of patients will develop an aggressive disease, either early progression (< 2 years) or tFL, at a rate for the latter occurrence estimated at around 2-3% per year (***Figure 8***). When FL is diagnosed in its early stages (I and II), radiotherapy may be considered for curative purposes. However, FL is still considered as an uncurable disease, each relapse requiring adapted therapy (cf. paragraph II.1.5). The advent of anti-CD20 monoclonal antibodies has dramatically improved management of the disease and affected patients have a median OS (overall survival) that can span up to 10 years with a minor impact on quality of life [186,187].

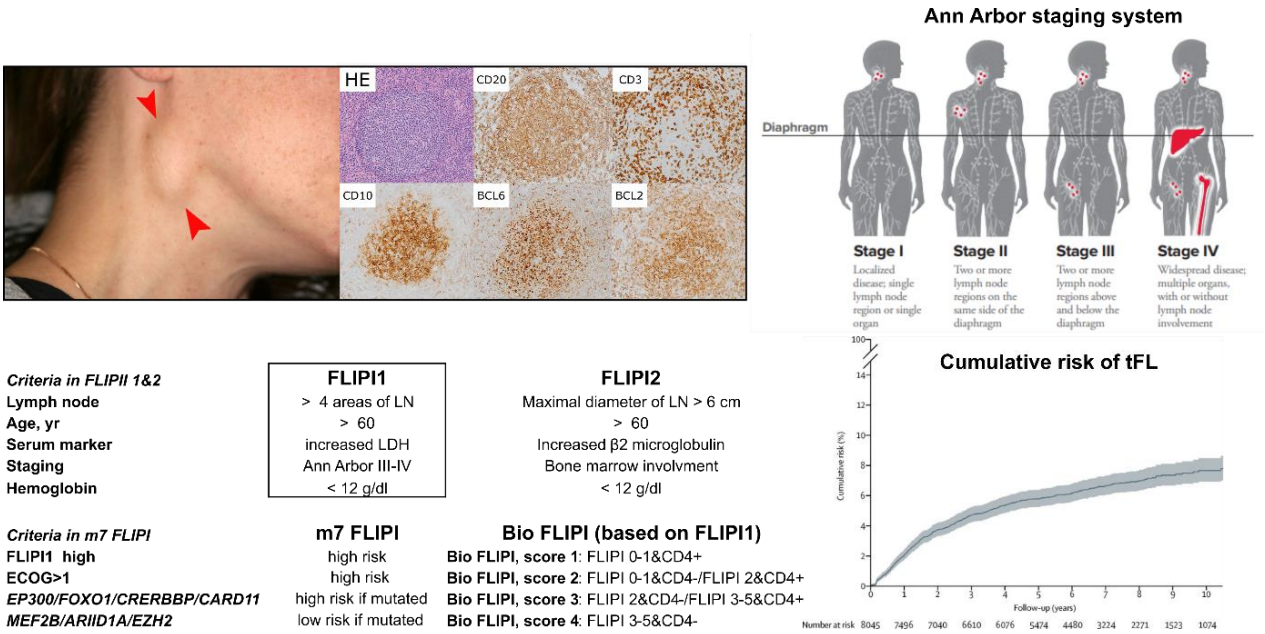


Fig 8: Follicular lymphoma at a glance

Superficial lymphadenopathy & main histology features of FL, adapted from Choi et al (2018) [188]; Ann Arbor Staging system (Leukemia & lymphoma society); FLIPI scores (FLIPI1 remains the most used in routine); cumulative risk of tFL, adapted from a meta-analysis of >10000 cases, in Federico et al (2018) [189]

II.1.3 Diagnosis and disease assessment

Histological examination is the gold standard for FL diagnosis. The best type of specimen, in order to obtain the most accurate diagnosis, remains an excisional lymph node biopsy whenever possible. A needle biopsy should be avoided but unfortunately remains too often the type of material provided to pathologists. Histological assessment of hematoxylin and eosin stained sections show replacement of normal lymphoid follicles by vaguely nodular or diffuse proliferation of small or medium-sized cleaved B cells (centrocytes) and larger non-cleaved B cells (centroblasts). The number of centroblasts per high power field allows the establishment of histological grade according to the WHO (grade 1-2 tumors contain less than 15 centroblasts, grade 3A more than 15 with remaining centrocytes, grade 3B more than 15 with no remaining centrocytes). Grade 3A with diffuse growth pattern and grade 3B are to be classified as DLBCL [172]. In addition to B cells, tumor microenvironment is composed of a variable amount of T cells, follicular dendritic cells and histocytes. Immunophenotype of FL cells helps to confirm diagnosis: typically, tumor cells are CD20+, CD3-, CD5-, BCL2+, BCL6+, CD10+, CD23-, CyclinD1- (***Figure 8***). Ki67 is usually low to moderate in grade 1-2 tumors and seems to correlate with grade disease. Immunophenotype of FL cells is rather similar with flow

cytometry (CD20+, CD19+, CD5-, CD10+, CD23-), expression of either immunoglobulin κ-chain or λ-chain confirms monotypia. Ultimately, additional molecular testing may be helpful for the pathologist when facing an ambiguous lesion: *IgH-Bcl2* translocation t(14; 18) assessed by fluorescence in situ hybridization (FISH) on FFPE section favors FL diagnosis, clonality study of the *IgH* gene/BCR variable region (using BIOMED 2 procedure) confirms the monoclonal nature of the population [190].

Standard initial staging evaluation requires diagnostic imaging with computed tomography (CT) or a fluorodeoxyglucose positron emission tomography (FDG-PET) scan when available in order to establish a complete mapping of involved sites (nodal, extra-nodal, visceral). Imaging is mandatory to assign stage I/II [191]. Calculation of total metabolic tumor volume (TMTV) raises increasing interest as a tool for prognostication (both at the time of diagnosis and for end of treatment response assessment), despite not being current standard practice [192]. Finally, bone marrow examination is part of the initial FL staging list.

II.1.4 Prognosis and risk stratification

The overall survival (OS) rate for FL at five years is 89% according to SEER database and median survival is approximately 10-12 years, 5 years in case of transformation [173,193]. Among the numerous clinical and biological models available for physicians to establish long term outcome, the Follicular Lymphoma International Prognosis Index (FLIPI) appears to be the most widely used and validated tool. It was established in 2004, based on data generated from > 4000 patients before the advent of Rituximab, and combines clinical and laboratory findings, allowing the stratification of patients into low, intermediate or high-risk categories of decreased OS, the latter corresponding to refractory disease, early progression or tFL (***Figure 8***) [194]. The FLIPI2 score, published in 2009, is slightly different and incorporates tumor bulk, bone marrow violation and β2 microglobulin [195]. FLIPI2 was developed using parameters more reflective of treatment in Rituximab era and seems to predict PFS more accurately than FLIPI1, even though this superiority remains inconsistently reported. In 2015, m7-FLIPI, sometimes called molecular FLIPI, proposed, in addition to classical FLIPI criteria, to incorporate mutational status of 7 frequently mutated genes in FL (*EP300, CREBBP, CARD11, MEF2B, EZH2, ARID1A* and *FOXO1*). M7-FLIPI, despite its limited use routinely to date, helped reassigning half of the FLIPI-classified high-risk patients as low-risk of progression [196]. Very recently, BioFLIPI incorporated the lack of CD4 intrafollicular T cells in FLIPI1 as a high-risk factor, underpinning the increasing evidence of the critical role played by microenvironment [197]. The Groupe d'Etude des Lymphomes de l'Adulte (GELA) developed the Primary Rituximab and

Maintenance (PRIMA) prognostic index (PRIMA-PI) based on the results of the PRIMA trial that involved 1135 randomized patients after chemoimmunotherapy and allowed stratification into low-, intermediate- or high- risk PFS according to only two criteria: bone marrow involvement and β2 microglobulin [198]. Huet *et al* further established a two-tier disease progression score based on the expression level of 23 genes that predicted PFS independently from FLIPI [199]. In another study, circulating tumor and cell free DNA also correlated to PET-TMTV and PFS [200]. In parallel with models built to estimate prognosis at the time of diagnosis, early progression within 24 months of treatment with chemoimmunotherapy (POD24) proved to be a robust prognostic tool since patients with early progression have a poorer outcome than the others [201].

Overall, in spite of their increasing accuracy and sometimes informative dimension about FL biology (m7FLIPI or BioFLIPI for instance), the use of these prognostic tools in FL remains limited since they do not provide information about how to best treat patients. Indeed, the choice of appropriate therapy is highly individualized and risk stratification models lack sensitivity and specificity to determine treatment decisions accurately.

II.1.5 Therapeutic management

FL is thus heterogeneous and the choice of first-line therapy is polyfactorial. Stratification is usually determined by disease stage.

Localized/limited stage disease (approximatively 10-15% of the patients) can be potentially cured with the use of radiotherapy (24 Gy involved field radiotherapy administered in 12 fractions) [202]. In cases when radiotherapy cannot be used safely, or in patients with mildly symptomatic disease, anti-CD20 monoclonal antibody Rituximab monotherapy can be proposed [203].

Not all patients diagnosed at advanced stage require immediate therapy. It depends on their clinical presentation or their will to receive treatment and its impact on quality of life. Indeed, the *watch and wait* approach is routinely favored for patients with no symptoms or low-tumor burden, while rapidly growing tumor burden, symptomatic disease or significant tumor burden impose induction chemotherapy. Most common schemes rely on CHOP (Cyclophosphamide, Doxorubicin hydroxide, Vincristine sulfate, Prednisone) in addition to an anti-CD20 monoclonal antibody, either Rituximab (so called R-CHOP) or Obinutuzumab [204,205]. Alkylator Bendamustine containing backbone seems to be as effective as R-CHOP with better tolerated side effects. However, clinical practices remain variable among the haemato-oncology community [206]. The RELEVANCE trial showed similar efficiency of Rituximab plus

Lenalidomide (an immunomodulatory agent) compared to R-CHOP/R-CVP (Cyclophosphamide, Vincristine sulfate, Prednisone) as an induction therapy, but this combination appeared slightly less well tolerated [207].

To date, a two year maintenance treatment with Rituximab improves PFS (but with no proven impact on OS) with a better tolerance than any other maintenance options [208,209]. Given the natural history of FL, which is characterized by multiple recurrence, additional therapeutic lines will be required throughout the course of this disease. The choice of treatment at the time of relapse is mostly based upon duration of remission after first line treatment. Thus, observation remains an acceptable attitude for relapsing patients (with imaging and/or pathology proven relapse) with no symptom nor high tumor burden. Symptomatic localized cases with low tumor burden may receive radiation monotherapy or Rituximab monotherapy. Most of the time, after long duration of remission, options for induction are chemoimmunotherapy in addition to Rituximab maintenance, yet, Obinutuzumab plus Bendamustine followed by Obinutuzumab maintenance showed efficiency in patients who did not receive any of these agents in prior therapy [210]. Recently, PI3K inhibitors (Copanlisib>Idelalisib, Duvelisib) have recently proved to be efficient with various toxicity profiles and are licensed with relapsing patients who received >2 prior lines of systemic therapy [211–213].

Early relapsing patients run a high risk of a poor outcome and represent a therapeutic challenge. Some of them might be eligible for autologous stem cell transplantation, but limited prospective data is available in regard to this treatment [214]. Lenalidomide–Rituximab combination has proven its efficiency following short remission [215]. In all cases and whenever possible, non-cross therapy is advised in relapsing patients followed by Rituximab maintenance.

PI3K inhibitors are within the scope of clinical trials for patients with POD24 and, more generally, evaluation of emerging therapies is a growing field of investigation for relapsed/refractory FL patients. We can mention BCR pathways inhibitor Umbralisib, anti-PD1 inhibitor Atezolizumab, or anti-CD47 therapies Hu5F9-G4 [216–218]. Finally, it is worth reporting biphenotypic therapy (anti-CD20/CD3 Mosunetuzumab) or the rising interest in reprogrammed chimeric antigen receptor (CAR)-T cell (for example, anti-CD19 CAR-T cell therapy Tisagenlecleucel) that can lead to long-term remission but which is still in its early stage of development [219,220]. Finally, the high frequency of mutations found in chromatin modifiers (cf. paragraph II.2.2.2) represents a rapidly emerging strategy for drug development in FL and histone deacetylases inhibitors such as Abexinostat for instance are being assessed in phase II trials [221]. The place of EZH2 inhibitors in therapeutic armamentarium for FL will be detailed in the 4th paragraph of this section.

II.2 Follicular lymphoma biology

Overt FL is considered as resulting from the accumulation of multiple genomic alterations through time. In spite of the similarities that exist at multiple levels between FL cells and mature GC B cells (centrocytes and centroblasts), lymphomagenesis of FL is a process which is believed to start early in B cell ontogeny and, from a clinical point of view, decades before the onset of symptomatic disease. This process has been extensively reviewed [222–224]. In this paragraph, I will present the most recurrent and persistent key events known to drive FL tumorigenesis, most of which have been identified and confirmed through several high-throughput sequencing based studies. I will ultimately summarize reported interactions between FL cells and their microenvironment (ME).

Before going deeper in FL biology *per se*, germinal center normal biology is briefly outlined in ***Figure 9***.

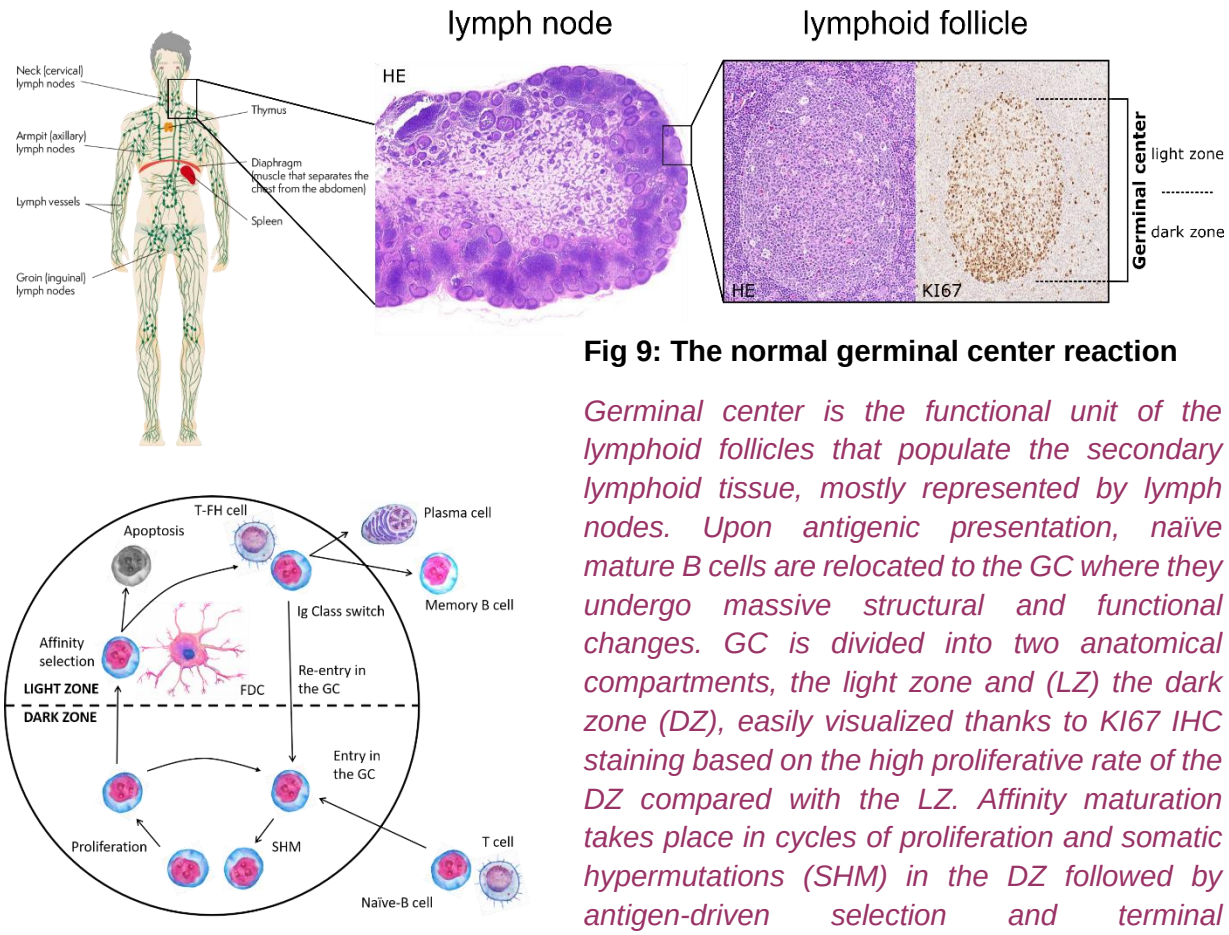


Fig 9: The normal germinal center reaction

Germinal center is the functional unit of the lymphoid follicles that populate the secondary lymphoid tissue, mostly represented by lymph nodes. Upon antigenic presentation, naïve mature B cells are relocated to the GC where they undergo massive structural and functional changes. GC is divided into two anatomical compartments, the light zone and (LZ) the dark zone (DZ), easily visualized thanks to KI67 IHC staining based on the high proliferative rate of the DZ compared with the LZ. Affinity maturation takes place in cycles of proliferation and somatic hypermutations (SHM) in the DZ followed by antigen-driven selection and terminal differentiation in the LZ, in close collaboration with the micro environment. The succession of events happening in the GC is called the GC reaction.

II.2.1 Genesis, precursor lesions

Reciprocal translocation t(14;18) (q32;q21) is often referred to as the genetic hallmark and quintessential event observed in FL cells. Indeed, it is detected in up to 90% of cases and occurs at the pre-B cell stage. Reports and studies about t(14;18) are numerous [225–230]: *Bcl2* gene is placed under the control of transcriptional regulatory elements of *IgH* gene, which results in the constitutive expression of anti-apoptotic protein BCL2. BCL2, in turn, inactivates pro-apoptotic proteins (BAX, BAK, BIM, PUMA). Translocation (14;18) stems from the constitutive fragility at CpG sites upstream *Bcl2,* leading to breaks and recombination with *IgH* during the VDJ recombination. Low rates of t(14;18) carrying cells (thereafter t(14;18) + cells) circulate in the peripheral blood of up to 70% of healthy individuals (0.1-10 cells/million) who will never develop overt FL. Therefore, t(14;18) alone is not sufficient to fully promote malignant transformation. However, a frequency of >1 cell/million is associated with a 23-fold increased risk of FL [231–233]. Some of the t(14;18)+ cells form restricted clonal populations characterized by functional memory B cell markers (IgD+/CD27+), suggesting that t(14;18)+ cells have been able to survive and differentiate through the GC reaction. These cells are sometimes referred to as FL-like cells (FLLC) (***Figure 10***). In a mouse model, FLLC have been shown to undergo iterative cycling in the GC and subsequent exposure to activation-induced cytidine deaminase (AID) before eventually experiencing progression to advanced precursor stages of FL [234]. This observation strengthens the concept of long-living FL functional precursors giving rise, upon acquisition of additional molecular alterations (the most frequent ones are detailed below), to committed FL cells. The concept of FLLC-driven pathogenesis is further supported by the observation of concomitant FL in both a donor and recipient years after allogenic transplantation, both tumors harboring identical *Bcl2* breakpoints [235].

Finally, t(14;18)+ cells can be incidentally found in lymph node tissue sections in apparently normal/reactive lymphoid follicles: yet immunohistochemical studies show that the immunophenotype of the cells that reside in these follicles is Bcl2+/CD10+. Such a finding defines the WHO recognized entity of *in situ follicular neoplasia* (ISFN), which is associated with a risk of developing FL at around 5% and might, consequently, represent a precursor of classical so-called FL. Similar IHC findings in a lymphoid follicle with partial architectural alteration correspond to a more advanced precursor of FL called *partial involvement by FL* (PIFL) [172].

Additional oncogenic hits are further required for t(14;18)+ cells to achieve complete progression toward FL. If the sequential acquisition of these hits has been long accepted as the global scheme of FL pathogenesis, some of the canonical mutations described in FL have been shown to arise early in B cell development in a contemporary way with, or even before

the acquisition of t(14;18). Those early B cells encompass multiple genetic alterations that are found to be present throughout evolution of the disease and are called *common progenitor cells* or *earliest inferable progenitors* (CPC or EIP) [236–238]. The sequence of acquisition of molecular events in CPC leading to overt FL remains partly mysterious, nevertheless, there is robust evidence that CPC display altered phenotype observed in FL once entering the GC reaction.

Indeed, upon antigenic stimulation in the peripheral blood or tissue, mature naïve CPC are relocated to the GC and undergo multiple rounds of division along with AID driven somatic hypermutations in the highly toxic and DNA damaging dark zone of the GC. However, the cells maintain physiologic functions such as functional IgM, and ongoing Bcl6 or AID activity. Eventually, some of them can reach memory B cell stage and further re-enter the GC reaction multiple times and undergo clonal expansion after a second antigen encounter.

Ultimately, CPC are very likely to survive throughout disease evolution in “reservoir niches” (bone marrow or elsewhere), resist standard therapies and be responsible for multiple relapses or transformation.

Thus, FL genesis can be depicted as a clonal dynamic between various GC B stages rather than a cell simply frozen at the GC stage, as recapitulated in ***figure 10*** [230].

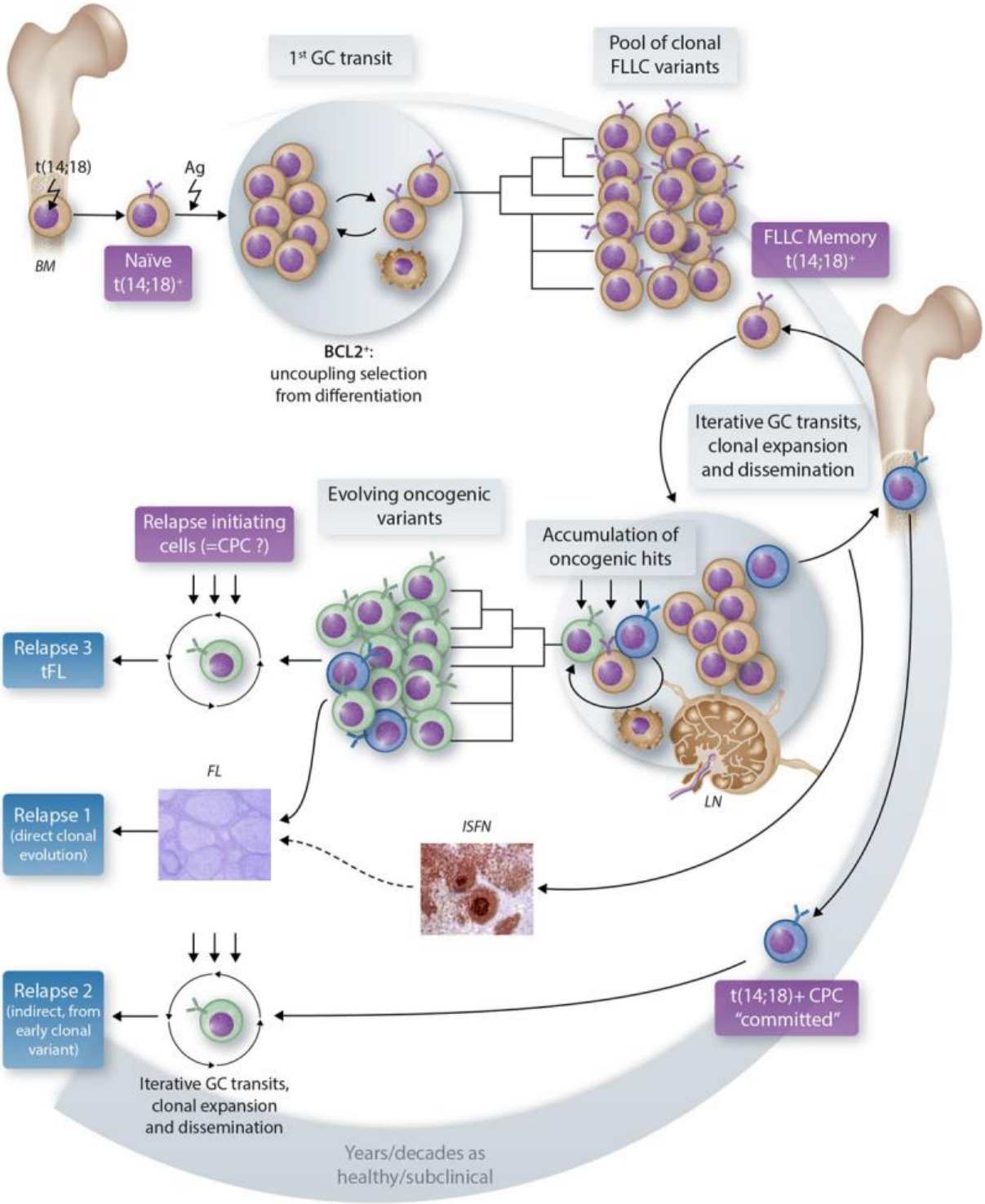


Fig 10: Outline of current model for follicular lymphoma pathogenesis, *from Brisou et al (2021)* [230]

II.2.2 Genomic, epigenomic alterations defining CPC and committed follicular lymphoma cells

II.2.2.1 Genome wide alterations

Only a few major studies depict the cytogenetic landscape of FL. 160/165 FL harbored secondary chromosomal alterations in addition to t(14;18) a study by Horsman *et al* up to 8 additional changes/cases were reported. The most recurrent changes (>10% of the cases) included +X, 1q21-44, +7, +12q, +18+, del 1p36, del 6q, del10. However, the number of secondary large-scale events was highly variable among the cohort and no recurrent combinatory pattern of alterations could be found [239]. Almost two decades later, Kriedel *et al* refined these findings retrieving up to 50 large scale alterations per sample (in addition with SNV found at a rate of 4000-1000 and 200-800 indels/cases), and a high frequency of loss of heterozygosity (LOH). FL still remains considered as a quite genetically stable malignancy [240].

Interestingly, genetic imbalance represents for some tumor suppressor genes a first oncogenic hit. A representative example in FL is the deletion of 1p36 that harbors *TNFRSF14* locus, along with a frequent loss-of-function mutation found in the remaining allele (cf. paragraph below). Similarly, a recurrent gain in chromosome 12 has been associated with impaired cell cycle control *via* CDK4 increased activity, which is relevant to FL tumorigenesis, given the exuberant cycling phenotype of FL cells [241].

In sum, it is very likely that upcoming whole genome sequencing coupled with CNV analyses will help decipher alterative mechanisms crucial for lymphomagenesis.

II.2.2.2 Epigenetic alterations

The most striking finding since the advent of massive parallel sequencing of FL series, which took off from around 2010, is the discovery of a multitude of recurrently mutated genes involved in chromatin modification (chromatin modifying genes, CMG), mostly histone modification. The notably high frequency of these alterations, along with modifications of DNA methylation and mutations in genes involved in chromatin remodeling, led the scientific community to consider FL (and also GCB-DLBCL) as a malignancy addicted to epigenetic perturbations. The section below summarizes the main traits of these recurrent alterations. Some alternate, yet recurrent mutational events will eventually be described, specifically when literature reports noticeable links with *EZH2*, in regard to the scope of this thesis.

Chromatin modifying genes

Pioneer studies form Morin *et al* in 2010 and 2011 reported the remarkably high frequency of mutations in chromatin modifying genes (CMG) in FL [242,243] (see also Okosun *et al* and Lunning *et al* ***Figure 11***). Since then, they have emerged as a keystone in FL biology. Unlike myeloid malignancies and T cell lymphoma, which exhibit mutations in DNA methylation modifiers, FL and GC-derived B cell lymphoma in general are characterized by mutations in genes catalyzing post-translational modifications of histones. It is interesting to note that mutations in CMG are strikingly more frequent in FL than in GCB-DLBCL. It is possible that this reflects the closer biology of FL to normal GC compared to DLBCL, underlying, therefore, the crucial role played by CMG in GC physiology. Moreover, more than 70% of FL harbor at least two mutations in CMG, with evidence of a model of serially acquired mutations through tumor progression. This pattern may relate to the increasing selective advantage conferred to FL cells by chromatin alterations through time, as a prototypical Darwinian trait for FL.

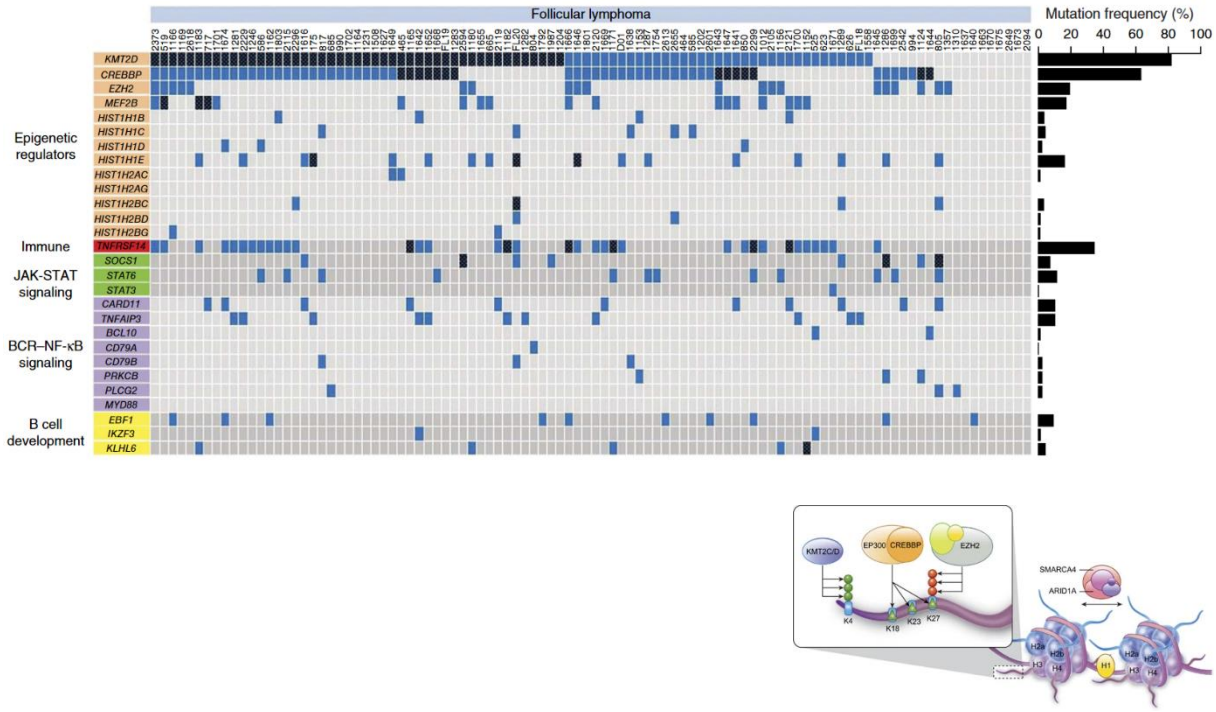


Fig 11: Most frequently mutated genes in FL, *adapted from Okosun et al (2014) and Lunning et al (2015)* [244]

KMT2D (a.k.a. *MLL2*) is by far the most frequently mutated gene in FL (up to 89% in the 2nd study by Morin *et al*, where 31/35 fully sequenced *KMT2D* loci were mutated with 18 truncating, 22 indel, 4 splice site, 3 SNV, mutations and with 15/35 harboring at least 2 mutations, in trans, leading to a virtually complete loss-of-function of the protein). As its name suggests, *KMT2D* is a SET domain containing lysine methyltransferase involved in H3K4 monomethylation, at

enhancers mostly [245]. Mutations reported in FL are mostly heterozygous but rare cases are homozygous, along with copy number neutral LOH (cnLOH). Notably, other MLL family member *KMT2C* is mutated in up to 15% of FL in a non-mutually exclusive manner with *KMT2D* suggesting, at least partially, non-redundant functions between these two enzymes. Both *KMT2D* and *KMT2C* can act alone or also interplay within the COMPASS complex [246]. The vast majority of *KMT2D* mutations are thus inactivating mutations, either through nonsense or frameshift mutations, or through mutations of the C ter region of the SET domain, leading to a reduced enzymatic activity. *Kmt2d* loss has been shown to induce a defect in B cell maturation along with a reduction of class switch recombination, in parallel with a reduction of H3K4me1 at GC B cell specific enhancers [247]. As for *EZH2* mutations, protumorigenic role of *Kmt2d* in mice models has been demonstrated in the context of Bcl2 rearranged cells [248]. The time of acquisition of *KMT2D* mutation through natural history of FL cells remains undetermined.

Similar cooperation with Bcl2 overexpression has been highlighted in the case of *Crebbp* mutation in mice [249]. Indeed, *CREBBP* is the second most frequently CMG mutated in FL (up to 70%). CRREBBP is a lysine acetyltransferase that acetylates notably H3K18 and H3K27 as well as non-histone proteins. As for *KMT2D*, *CREBBP* enhances transcription and the majority of mutations found in FL are missense mutations that target the lysine acetyl transferase (KAT) domain of the protein leading to a reduced acetylation [250]. Mutations of *CREBBP* are thus loss-of-function mutations arguing for a tumor suppressive role in FL biology. Paradoxically, the mutational rate of *CREBBP* is higher in low-grade tumors but appears to be associated with a worse prognosis in the m7FLIPI model. *CREBBP* homologous *EP300* weight on prognosis is even heavier in m7FLIPI [196]. One remarkable finding about mutant *CREBBP* in Green *et al* thorough phylogenetic study is that it happened to be the only recurrent mutation present at the very early stage (early inferable progenitors) along with t(14;18), in a subset of cells responsible for not only overt FL but also disease relapse [237]. Kridel *et al* analysis of comparative mutational landscapes from FL at diagnosis, in early progressors (POD24) and tFL, confirms the recurrence of *CREBBP* mutation throughout disease evolution [240]. From a more functional point of view, H3K27ac depleted loci upon *CREBBP* mutation happened to be enriched in Bcl6, a transcriptional repressor often overexpressed in FL that plays a crucial role in GC reaction orchestration and maintenance, underpinning the interlaced role played by these two frequent alterations in FL (cf. below) [250]. Interestingly, a reduction in expression of MHC class II genes along with a loss of CREBBP at enhancers of these genes and reduced tumor infiltrating T cells transcriptomic signature were observed in mutant cells [237]. Indeed, *CREBBP* mutation decreased stimulation of CD4 T cells, these findings ultimately converging for a role of *CREBBBP* mutation not only played *via* differentiation blockade but also supportive of immune evasion. Of note, non-histone protein acetylation by CREBBP, and notably p53,

has shown a potential role in lymphomagenesis by the mean of reduced DNA damage response [238].

Along with *CREBBP* and *KMT2D* loss-of-function mutations, *EZH2* is the third most frequently CMG mutated in FL, but unlike the previously mentioned genes, *EZH2* mutations have been extensively reported as gain-of-function mutations, on the one hand, and, on the other hand, might be sub clonal at diagnosis, sometimes acquired during disease evolution. A more detailed review of literature about current knowledge about *EZH2* mutant FL will be provided in the 3rd. paragraph of this chapter.

Overall, there is strong evidence that *KMT2D*, *CREBBP*, and *EZH2* mutations share a common function in controlling GC B cell differentiation programs.

Among other CMG altered in FL, it is relevant to mention the existence, even sparse in FL contrary to what is observed in Burkitt lymphoma (another subtype of GC derived B cell lymphoma), of inactivating mutations of members of the SWI/SNF chromatin remodeler complex (e.g. *ARID1A, BCL7A*) [251]. Such alterations are proposed to intensify reliance on activity of *EZH2*, antagonist of SWI/SNF activity. Both disruption of the PRC2 complex and inhibition of the *EZH2* SET domain have been demonstrated to promote cell death in a SWI/SNF deficient background. Consistently, tumors carrying SWI/SNF inactivating mutations are particularly sensitive to genetic and pharmacological inhibition of PRC2. This has been demonstrated in epithelioid sarcomas, where patients with mutations in the SWI/SNF complex member SMARCB1 respond better to EZH2 inhibitors [252]. It might thus be meaningful to mention that *ARID1A* mutations confer a better prognosis at diagnosis, as *EZH2* mutations do, in m7FLIPI [196].

DNA methylation

Mutations affecting gene coding for DNMTs are a rare event in FL, yet DNA methylation pattern is altered. FL cells are characterized by genome-wide hypomethylation contrasting with hypermethylated DNA state at Polycomb target genes and loci repressed by PcG marks in ESCs [253]. Moreover, tumor suppressor genes, such as cyclin dependent kinase inhibitors (p15, p16, p57), are found to be hypermethylated thus repressed, mostly at the time of transformation [254]. Indeed, methylation profile has been suggested to foster FL precursor cells to undergo fully developed FL, and primary FL to tFL. Some authors suggested that a cooperation between histone modification and methylation would lead to a differentiation blockade [255].

II.2.2.3 Alternate key pathways altered in FL

Bcl6 expression dysregulation

Bcl6 is a transcriptional repressor whose expression plays a central role in GC formation and maintenance. Schematically, down regulation of Bcl6 allows exit of GC and terminal differentiation [256]. Given its pivotal role in GC orchestration, aberrant expression of Bcl6 has a leading impact in FL. Bcl6 is thus overexpressed in >95% of FL through various mechanisms (chromosomal translocations leading to constitutive activation, AID induced activating mutations [256], indirect activation *via* mutations of Bcl6 transcriptional activator *MEF2B* [257], reduced BCR signaling mediated by NOTCH2 pathway repression [258], *CREBBP/EP300* mutations[250]). Bcl6 recruits and stabilizes several factors to chromatin (including HDAC or HDMT) in order to silence specific signaling pathways, thus leading, among others, in enhanced tolerance to DNA damage. Such tolerance is crucial in the GC and eventually facilitates hypersomatic mutations of IgG. Bcl6 has also been proposed to cooperate with *EZH2* in order to recruit non-canonical PRC1-BCOR-CBX8 complex and, in turn, repress genes involved in the exit of the GC, thus favoring lymphomagenesis [259].

Activation of mTORC1 signaling pathway

Recurrent activating mutations tend to co-occur in *RRAGC* and/or vacuolar H+ adenosine triphosphate ATPase coding genes *ATP6V1B2* and *ATP6AP1*, which both activate mTORC1, reflecting abusive activation of the mTOR signaling pathway [260]. Furthermore, mTORC1 is repressed by SESTRIN1, which appears to be targeted by 6q deletion in FL. Interestingly, it has been shown that EZH2 Y646 mutant specifically represses SESTRIN1, resulting in enhanced mTORC1 activation in comparison with the WT enzyme. Of note, either deletion of *SESTRIN1*, *EZH2* mutation or *RRAGC* mutation tend to be mutually exclusive, pointing to alternate pathways for mTORC1 signaling activation. This mechanism is further confirmed since *SESTRIN1* deletion mitigates EZH2 inhibition response in *EZH2* mutant but not WT cells. Moreover, *EZH2* mutant cells are sensitive to mTORC inhibitor unlike WT cells [261]. Whether these observations open the door for potential therapeutic strategies in *EZH2* mutant FL remains speculative but are nonetheless interesting as they illustrate the cooperation or the complementarity (or maybe coevolution?) of multiple mechanisms in FL precursors in order to promote cancerous process.

Disruption of TNFRSF14-BTLA biding

As mentioned above, *TNFRFS14* (a.k.a. *HVEM*) is another example of a tumor suppressor gene frequently altered in FL, either *via* inactivating mutations or *via* loss of chromosome 1p [262]. Both losses and mutations lead to a reduction in TNFRFSF14 expression with altered binding with BTLA [263,264]. *Tnfrsf14* KO fosters lymphomagenesis in Vavp-Bcl2 mice [265] due to enhanced immune evasion. Indeed, TNFRSF14-BTLA axis not only stimulates BCR signaling but also increases interactions with surrounding cells, such as FDCs and fibroblasts, as well as follicular helper T cells (Tfh) through secretion of cytokines. Interestingly, both *TNRSF14* and *EZH2* mutations have been reported to co-occur at the time of transformation from FL to tFL in some cases. Although purely observational, this study actually revealed the high mutational frequency of *TNFRSF14* together with recurrent gains of *IL4-R* as well as *EZH2* mutations in a series of paired FL/tFL, which could suggest a link between B & T cells interactions in the microenvironment at the time of transformation [266]. Of note, a significant co-occurrence between *EZH2* and *TNRFSF14* found in GCB-DLBCL echoes these findings [243].

Other recurrent mutations

Not all the numerous recurrent mutations found in FL could be reported here but one should keep in mind the recurrence of genes involved in JAK-STAT signaling pathways such as activating mutations of *STAT6* or inactivating mutations of JAK/STAT inhibitor *SOCS1* that are seen in up to 11% and 25% of FL cases respectively [267,268].

Li *et al* & Okosun *et al* almost simultaneously reported the high frequency of nonsense mutations found in histone genes with an overall rate of 27% and 28% out of 122 and 132 cases respectively harboring a mutation in at least one histone H1 gene [269,270]. These loss-of-function mutations are very likely to alter histone binding to chromatin and thus expression of substantial number of genes.

Interestingly, not all mutations in FL emphasize functions classically associated with GC B cell biology but, instead, target functions relevant to a later B cell stage, such as NF-κB signaling or BCR signaling (*CARD11* or *MYD88* mutations, for instance). Indeed, these mutations are preponderant ABC-DLBCL (activated B cell- DLBCL) [270].

II.2.2.4 Follicular lymphoma microenvironment

As in all B cell lymphoma subtype, FL cells orchestrate a microenvironment (ME) made of inflammatory and mostly immune and stromal cells that sustain not only tumor cell growth, but

also favor immune evasion (as recapitulated by Devan *et al*, ***Figure 12***). Based on the similarities existing between FL and GC, at both cellular and architectural levels, it is not surprising that FL cell growth relies, to a certain extent, on a pre-existing crosstalk that is pivotal in the normal GC reaction. This dependency on ME also possibly explains why FL cell lines have been difficult to establish. Germinal center 3D organoids therefore represent a promising model for future assays [271]. The importance of ME is underpinned by the prognostic value of gene expression signature derived from ME cells for patient outcome [272,273]. Interactions between FL cells and ME are thus most likely to result from a hijacking of the GC ME physiological functions. As for normal GC B cells, FL cells in a lymph node are strongly dependent on Tfh and FDCs, but also seem to abnormally upregulate regulatory - cell (Treg) population [274]. When localized to the bone marrow, FL ME happened to be enriched in mesenchymal stromal cells and myeloid cells. ME cells not only interact with FL cells but also engage in a constant crosstalk with each other, enabling tumor growth. Features of the ME of FL are abundantly described in the literature [275,276].

Strong evidence supports the secretion of IL4 and IL21 by Tfh cells, which in turn activate survival pathways in FL cells. High production of IL4 by Tfh in FL has been shown to be induced, as for the chemokine CXCL12 secretion, by both FDCs and fibroblastic reticular cells [277,278]. Of note, the normal function of CXCL12 is to polarize the GC into dark and light zones - one example of how FL cells hijack normal GC function. Tfh also express CD40L, ligand of CD40 that is expressed on the surface of FL cells. CD40-CD40L creates a feedback that stimulates FL cell growth [279].

Another example of interplay between FL cells and the ME is the mannosylation of BCR by macrophages illustrating intervention from innate immune system in the ME. These added mannoses interact with lectins of the mannose receptor and with CD209 (a.k.a. DC-SIGN), expressed by macrophages in the microenvironment, which in turn drive FL cells to escape the normal BCR selection [280–282]. Beyond tumor growth and cell signaling pathways, ME thus plays a crucial role in immune evasion.

Impact on CMG mutations, specifically *CREBBP* and *EZH2,* in antigenic presentation repression is discussed elsewhere in this introduction.

In summary, several pathways are activated by FL cells to reprogram their ME and the concept of permissive immune FL niche is strongly supported by a multitude of studies. Development of innovative therapies leveraging ME biology in FL is an effervescent field of investigation as in many cancer types. Moreover, the strong interlace between FL cells and surrounding cells

might be testament of a malignancy that is the result of a co-evolutionary process of cells from various origins [236].

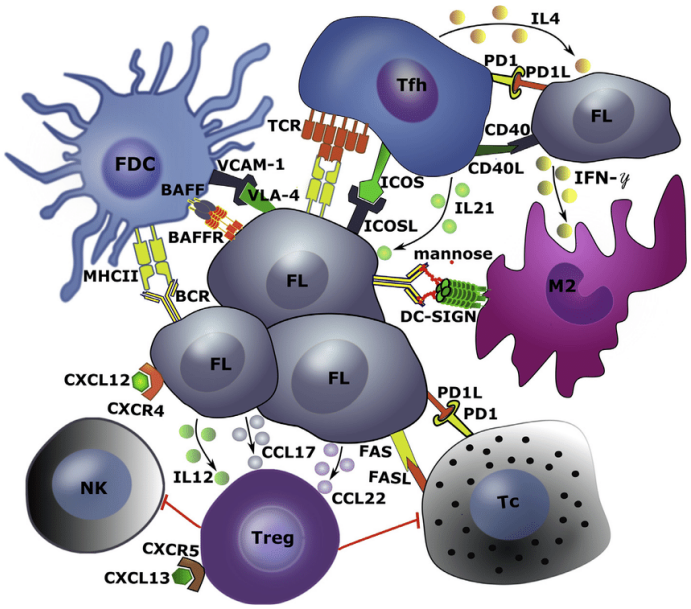


Fig 12 : Follicular lymphoma microenvironnement, *from Devan et al (2018)* [283]

II.2.2.5 Clonal architecture and evolutionary pattern of follicular lymphoma

From the first exome sequencing performed in tumor samples, high intratumoral clonal diversity appeared to be a striking characteristic of FL, with numerous clonal populations identified (inferred through allelic frequency assessment) being private/sub clonal [236]. Nonetheless, clonality in FL seems to follow a constant hierarchical trend with dominant/clonal events present throughout disease evolution. Exome sequencing coupled with copy number longitudinal studies further helped unravel certain traits of evolutionary patterns of FL progression through time and from FL to tFL [284,270,240,285].

Firstly, in a vast majority of cases, *IgH* locus sequence remains stable along disease evolution, in a clonal way, confirming the strong reliance on t(14; 18) for FL. Aside from t(14; 18), mutations present at a high clonal rate early in overt FL define what is often referred to as the CPC (or EIP) (cf. II.2.1) - and encompass altered CMG in a nearly mandatory way. To date, *CREBBP* mutations seem to be the most frequently enriched alterations in EIP, *KMT2D* mutations might appear slightly later in disease evolution [237]. Whether the mutation of *EZH2* is acquired early or late in FL natural history is debatable (cf. next section). In Green *et al* study, no mutation of *EZH2*was found at the level of EIP [236].

Evolutionary patterns of FL seem to split between a relapsing, progressive disease pattern on the one hand and transformed pattern on the other hand (tFL) [240]. In the former case, relapse samples are characterized by the presence of the same clones detected at the time of diagnosis with a progressive acquisition of minor clones in a linear fashion. The latter pattern is characterized by the proliferation of clones either undetectable or anecdotal at diagnosis. These clones are eventually selected and expand, sometimes with an extinction of the CPC (or at least undetectable in the main tumor burden). No single event appears responsible for transformation, which seems mostly to be the consequence of the accumulation of distinct and multiple alterations (such as mutation in *MYC*, *TP53*, or *B2M*). Both progression and transformation have been shown to be accompanied by an increase in the number of CNV along with an accumulation of mutations. Interestingly, Green *et al* demonstrated that mutations specific to relapsing tumors compared with diagnosis are prone to be motifs recognized by AID (or APOBEC), strengthening the model of multiple re-entries in the GC before relapse. Yet, high frequency of relapses suggest that CPC are resistant to standard therapies and there is a need in strategies to stratify treatment according to early genetic lesions identified in CPC.

II.3 *EZH2*^{Y646} mutations in follicular lymphoma, state of the art

One of the purposes of this thesis is to provide, through a combined mechanistic and translational approach, a better understanding of the role of recurrent point mutation *EZH2*^{Y646} reported in FL by R. Morin for the first time, a little over a decade ago. I will summarize the most relevant roles of EZH2 in normal B cells before reviewing the main characteristics of the mutant enzyme and its putative impact in FL biology.

II.3.1 Role of EZH2 in normal B cells

Given the highly restrictive nature of *EZH2*^{Y646} to GC B cell malignancies, an overview of the role played by *EZH2* in normal B cell provides relevant information for a clearer understanding of its role in FL. Of note, in regard to the uncertainty about the PRC2-independent functions for EZH2, I chose to focus on its canonical role.

II.3.1.1 Role of EZH2 in hematopoietic stem cells

EZH2 is critical in the B cell developmental process, and is particularly expressed at very early stages of B cell life, then subsequently, at the time of antigenic-induced activation and during the GC reaction with a crucial role in proliferation and Ig affinity maturation (cf. ***Figure 13***). In adult hematopoiesis, conditional inactivation of *EZH2* is dispensable for HSC function, most likely due to EZH1 compensation. Indeed, *Ezh1/2* double KO impairs HSCs function in a similar way as observed in *Suz12* or *Eed*-KO mice [286]. Deficient PRC2 impairs self-renewal of HSCs and leads to upregulation of PRC2 targets (including cell cycle inhibitors, thus mitigating self-renewal properties). Conversely, overexpressing EZH2 in HSCs was reported to enhance capacity for self-renewal and prevented HSCs from cell exhaustion. The interpretation of this study is, however, debatable as the functional consequences of overexpressing a subunit of the complex are unclear [287]. Other studies showed that heterozygous disruption of any PRC2 member (i.e. in PRC2 loss-of-function mutations) enhanced HSC activity, highlighting the importance in gene dosage at this stage and its impact in many hematological malignancies [288].

II.3.1.2 Role of EZH2 in early B cell development

EZH2 is highly expressed in lymphoid progenitors (pro- and pre- B stages) and it declines in resting mature B cells, till further major upregulation at the time of GC reaction initiation [289]. Impaired *EZH2* in B cell progenitors dramatically limits B cell compartment expansion [290]: EZH2 allows transition from pro- B to pre- B stages and is implicated in major events such as VDJ recombination of Igh, relating to, among other mechanisms, silencing of *Cdkn2a* [291,292]. EZH2 is further involved in the later part of the pre-B cell stage also through *Igk* recombination *via* its STAT5 mediated recruitment at *Igk* loci [293]. Of note, in the Souroullas *et al* study (detailed in the next section focusing on the role of *Ezh2*^{Y641} in lymphomagenesis), the mutation did not disturb the development of B cells in young mice [294].

II.3.1.3 Role of EZH2 in late differentiation: the GC reaction and plasma cell differentiation

During B cell activation, GC B cells (both centroblasts and centrocytes) upregulate and highly express *EZH2*. Once B cells exit the GC reaction, *EZH2* is finally downregulated.

Extensive work from Caganova *et al* and Béguelin *et al* have demonstrated that *EZH2* is necessary for GC formation along with activation and control of all the major steps of survival,

maturation and differentiation achieved by B cells in the GC [259,295–297]. By the mean of *CDKN2A* repression, *EZH2* controls the high proliferative rate of GC B cells, which, in turn, allows formation of high affinity antibodies. *EZH2* target genes in GC B cells partially overlap with *EZH2* target genes in embryonic stem and naïve B cells, which might reflect its role in preventing terminal differentiation through deposition of H3K27me3 at crucial targets. It includes genes involved in the exit from GC, such as *PRDM1*, *IRF4*, and *XBP1*, but also negative regulators of the cell cycle, such as *CDKN1A* and *CDKN1B.* Moreover, the loss of restoration of GC formation in the absence of EZH2 upon *Aicda* loss (coding gene for AID) provides evidence of the determinant impact of *EZH2* on AID activity, thus on somatic hypermutations.

Ultimately, EZH2 remains strongly upregulated in plasmablasts but is absent from plasma cells and quiescent memory B cells [298]. Indeed, at the time of plasma cell differentiation, EZH2 colocalize with Blimp1, fostering silencing of *Blimp1*-repressed genes [299]. The pivotal role of *EZH2* in terminal B cell differentiation is enlightened by the enhanced plasma cell formation upon EZH2 inhibition in plasmablasts *in vitro*.

The critical role played by *EZH2* in many stages of B cell late maturation might represent an option to investigate for EZH2 inhibitor as an immune modulator strategy in autoimmune conditions, such as lupus or asthma.

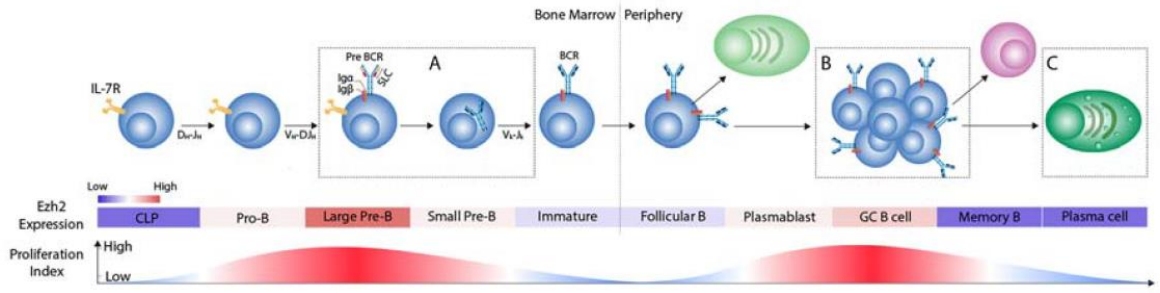


Fig 13: EZH2 expression level throughout B cell life, adapted from Nutt *et al* (2020) [300]

II.3.2 *EZH2* ^{Y646} mutation in follicular lymphoma

Y641 or Y646 ?

When first described in FL, EZH2 mutation was reported based on the nomenclature of* EZH2 *Refseq isoform C (NM_001203247), corresponding to the Y641 codon. Y641 also corresponds to the same alteration in Ezh2 according to mm10 annotated mouse genome. Slightly longer* EZH2 *Refseq isoform A (NM_004456) has now been widely used as the reference transcript to designate Y646 for the same residue. Thereafter, I will use Y646 to designate the mutation in human tumors and Y641 in mouse experiments in accordance with most widely used nomenclature in literature.

II.3.2.1 Discovery, main genomic and biochemical characteristics

In 2010, Morin *et al* were the first to report a rate of around 7% and 22% somatic mutations of *EZH2* (then Tyr641 at the time) in FL and GC-DLBCL respectively. These results were obtained upon both high throughput and Sanger sequencing of a cohort made of 221 FL and 320 DLBCL biopsies arising from both GC and ABC cell of origin, along with 80 other non-GC derived lymphoma subtypes, lymphoma cell lines and benign tissue samples. *EZH2* mutations were absent from every non-GC derived B cell lymphoma (except in 1/24 PMBCL, primary mediastinal B cell lymphoma), and were further confirmed to be somatic mutations based on the absence of the mutation in matching germline DNA from 9 mutated lymphoma patients. Of note, this description was an incidental discovery of a mutation in *EZH2* in one patient carrying a grade 1 FL with no associated t(14;18), for which deep sequencing of DNA and RNA was performed in order to better characterize this unusual simple karyotype-associated tumor [242].

Thereafter, several series confirmed this observation. The biggest series reporting *EZH2* mutations in FL to date (n=366) was published by Bödör *et al* in 2013 (***Figure 14***): Sanger sequencing of exons 16 and 18 resulted in 17% of *EZH2* mutant cases, however, NGS applied to the entire cohort allowed the identification of 43 additional mutations in 39 patients with a lower Variant Allele Frequency (VAF) than the mutations detected with the Sanger method, increasing the final mutant rate to 27% of the entire cohort [301]. *EZH2* mutations in FL never occur outside of exons 16 and 18. Out of the 148 cases analyzed with RNA-Seq from the PRIMA cohort, Huet *et al* reported 42 (28%) cases. Validation of these cases along with investigation of 11 extra cases using the Sanger method yielded to a total 46/159 (29%) mutant cases [302]. Thus, NGS presumably helps retrieve the highest rate of mutant cases compared with Sanger sequencing.

EZH2 mutation is always reported to be heterozygous in FL. Morin *et al* therefore raised the question of the allelic dosage in their 2010 publication: a significant preferential expression of the mutant transcript over the WT was noticed, but this series remained of limited effective size (4/13 *EZH2* mutant FL) while some other cases displayed a skewed expression toward the WT allele [243]. The equimolar amount of both WT and mutant alleles in lymphoma cells was eventually confirmed through precise mass-spec experiments, thus providing evidence that the effect of the mutant would not rely on a greater amount of one enzyme compared with the other one [88].

In a vast majority of the cases, these are missense mutations of tyrosine 646 (Y646) within the SET domain of *EZH2*. In Morin *et al*, out of the eight possible non-synonymous variants of Y646 codon, five were detected with a predominance of Y646F (49%) over Y646S (21%), Y646N (15%), Y646H (13%) and Y646C (2%). The nature of the single nucleotide changes

induced by these mutations indicate that they are very unlikely to result from AID-induced hypersomatic mutations, which specifically triggers C>T and G>A mutations [303]. This latter finding supports a pre-GC stage acquisition for *EZH2*^{Y646} mutations. Additionally, 1/25 of fully sequenced cases harbored another mutation in N635 (also affecting the SET domain), *in cis* with Y646 mutation. The same five variants of Y646 identified by Morin *et al* were present in Bödör *et al* and Huet *et al* series respectively, with a slightly different order in terms of frequency (Y646N: 36%/45%, Y646F: 27%/29%, Y646S:12%/16%, Y646H: 9%/5%, Y646C: 3%/5%). Of note, Bödör *et al* series revealed 4 multiple mutations (2 in the same allele, 2 in different reads) and 3 novel variants, all of them located within the SET domain in addition with *EZH2*^{Y646}: K634E, V637A and V679M. Huet *et al* series included four cases with A692 and one case with W629G mutation. Huet *et al* also investigated copy number and reported 23 cases (15%) with a gain of 7q36.1 at loci encompassing *EZH2*. No association was found between copy number and mutational status. Another series reported 24% of 12 FL cases with a gain or amplification of EZH2 at chromosome 7q [304].

E(z) (+), a mutation orthologous to mutant *EZH2* described in the *Drosophila* gene E(z), has been shown to induce a phenotype distinct from other loss-of-function mutations and was eventually reported as a gain-of-function [305]. However, Polycomb complexes containing E(z) (+) were not able to perform trimethylation H3K27 *in vitro* [306]. A similar observation was initially made in human setting: indeed, in Morin *et al* study, the 4 main mutant proteins (Y646F, S, N and H) were capable of associating and forming a PRC2 complex *in vitro* along with WT EED, SUZ12, RbAP48 and AEBP2 protein in Sf9 cells. Trimethylation activity of PRC2 assembled with the 4 mutants assessed through the use of cell-free biotinylated histone H3, with addition of SAM, appeared to be dramatically reduced compared to WT EZH2 (around 7 fold), leading to the conclusion that Y646 mutation conferred a loss of enzymatic function to EZH2, hence similar to the observations made with E(z) (+) [242] (***Figure 14***). However, Yap *et al* subsequently reported in B cell lymphoma cells and in *Ezh2* mutant transfected B cells from mice, with two different antibodies, an increased global level of H3K27me3, raising suspicion about the impact on PRC2 activity of the mutant [88].

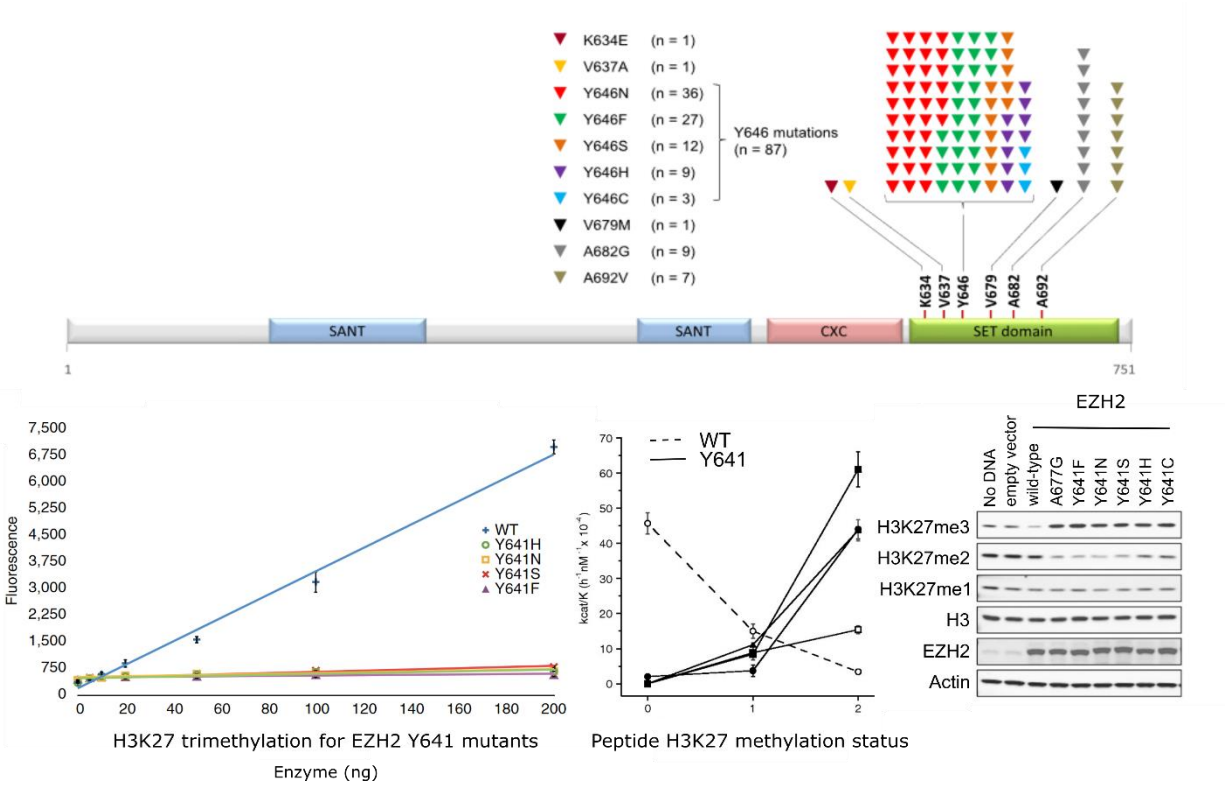


Fig 14: Enzymatic activity of mutant EZH2

Distribution of 87 EZH2 Y646 mutations out of 366 cases in Bödör et al (2013) (upper). Methyltransferase activity of each mutant and WT assessed independently in vitro by Morin et al (2010) showing a drastically reduced activity for the mutants (lower left), whereas Sneeringer et al showed the restricted ability of the mutant enzyme to catalyze the last steps of H3K27 methylation in comparison with WT (2010) (lower middle). Western blot of H3K27me level in MCF7 cells transiently transfected with various EZH2 mutant in McCabe et al (2012) (lower right)

Further biochemical characterization by Sneeringer *et al* showed that the enzymatic activity of WT EZH2 to perform 0-1-2 methyl substrate had a ratio of 9:6:1, whereas Y641 EZH2 had a ratio 1:2:13 and A677G EZH2 1.1:0.6:1 (***Figure 14***). Indeed, H3K27me3 level assessment through multiple lymphoma cells revealed the existence of one cell line, Pfeiffer, harboring *EZH2* mutation A677G in exon 18 (SET domain), with a high level of H3K27me3 similar to Y646 [87] (also shown by McCabe *et al*, ***Figure 14***). Overall, based on these enzymatic studies and the mandatory heterozygous fashion of EZH2^{Y646} in FL, it is widely accepted that the mutation confers an enzymatic gain-of-function through cooperation with the wild type allele. However, puzzling observations reported that the level of H3K27me3 was similarly increased in both homo and heterozygous isogenic mice with no compensatory effect of the paralogous enzymatic subunit of PRC2 *Ezh1* [294]. Moreover, given the ever cycling nature of FL cells, it is unlikely that the expression of *EZH1* compensate for impaired *EZH2* as shown by our lab [40].

Thus, the inherent ability for mutant enzyme EZH2 to catalyze H3K27me3 is not completely ruled out yet.

Subsequently, in regard to the description of the mutation, several groups sought to understand the consequences of the mutant at the structural level of the molecule. To this end, Yap *et al* conducted an analogical structural analysis of the SET domain of *EZH2* and proposed that the mutant led to a shift of product equilibrium with a dominant altered substrate catalytic specificity of the enzyme for its substrate. In other words, H3K27me2 would become the preferred substrate of the mutant enzyme (instead of mono or un methylated H3K27 for the WT enzyme). Consequences on chromatin would be a shift from di- to tri- methylation at PRC2 target loci [88].

McCabe *et al* also took advantage of a structural model to better characterize the substrate specificity of mutant EZH2 and showed that it participates in the orientation of the un- and monomethylated substrate and sterically restricts the enzymatic activity with demethylated substrate, leading to an increased activity towards a larger H3K27me2 substrate to rotate into position for a methyl transfer [307].

Crystallographic characterization of recombinant EZH2 SET domain further allowed the description of a model where the Y646 side chain would point inwards the catalytic locus, thus forming only hydrophobic and van der Waals interactions with surrounding protein atoms. This would occlude the lysine substrate binding channel, with the side chain hydroxyl forming a lone hydrogen bond with a water molecule. All these structural changes would ultimately bury the Y646 mutant in a hydrophobic cluster except for the solvent exposed tip of the phenyl ring, where the phenyl oxygen forms a hydrogen bond with a water molecule. Of note, regardless of the substituted amino acid encountered in the EZH2 mutant, this may result in a reduction in space occupied by the substituted amino acid, in turn enabling an increased productive binding of methylated substrate. Moreover, burial of EZH2 mutant Y646 in inactive conformation (i.e. free from EED and SUZ12) implies numerous broken internal contacts, thus altering the dynamics of the transition process between inactive and active state.

Some mutations localized outside of the active state reported in FL, such as A682, form van der Waal interactions with Y646 side chain, leading to indirect but similar effect on substrate specificity as Y646 mutation itself. In total, the mutation leads to a structural release of spatial constraints of the SET domain, thus facilitating the adjunction of a third methyl group to H3K27 [308,309]. Additionally, the crystal structures of the inhibitor-bound WT and Y646N-PRC2 were reported, underlying the important role played by a stretch of 17 residues in the N-terminal region of EZH2, in the stimulation of the enzyme activity, inhibitor recognition and then potential acquisition of a mutation-mediated drug resistance [310].

Finally, an elegant set of experiments from Sahasrabuddhe *et al* showed that EZH2 is stabilized by inhibition of proteasome pathway, in a β-TrCP ligase-dependent manner. Within this framework, EZH2 Y646 appears more stable and has a longer half-life than WT as it doesn't interact with β-TrCP and has reduced ubiquitination. This lack of interaction might stem from the fact that EZH2 Y646 is resistant to JAK2 mediated phosphorylation [311]. Therefore, even though the amount of both WT and mutant allele is equimolar in the nucleus, a transient quantitative superiority of the mutant cannot be ruled out based on these observations.

II.3.2.2 Mutational environment, phylogenesis and prognostic value of EZH2^{Y646}

Co-occurrence of *EZH2* mutations with other CMG is then the general rule in FL, which has mechanistic consequences. Indeed, there is a well-defined antagonistic relationship between H3K27ac and H3K27me3, suggesting that *CREBBP* mutations, which result in removal of H3K27ac, also impacts the deposition of H3K27me3 [312]. Moreover, *KMT2D* can also indirectly influence both H3K27ac and H3K27me3 through a COMPASS-like complex that recruits the UTX histone demethylating enzyme to enhancers, thereby facilitating the abrogation of H3K27me3 [246]. This indicates that loss of *KMT2D* may lead to the accumulation of H3K27me3 at enhancers where it is normally bound. Although many of the changes in H3K27me3 catalyzed by mutant EZH2 were observed at promoters and coding regions, a loss of KMT2D may allow mutant EZH2 to additionally silence enhancer elements.

As mentioned earlier, results are rather contradictory in regard to the clonal nature of *EZH2*^{Y646}: The longitudinal study from Bödör *et al* compared VAF of the *EZH2* mutant to VAF in other frequently mutated genes in 43 cases eventually enabling them to show that *EZH2* mutation was a clonal event in the vast majority of the cases (81%), which was confirmed by another study [284,301]. Nonetheless, 19% of the mutated cases appeared to display a sub clonal phylogeny and, in an anecdotal number of cases, *EZH2* showed a clonal phylogenetic pattern but with a very low VAF. In the longitudinal study by Green *et al*, however, 4/6 *EZH2* mutations (in a cohort of 22 cases) were found only in relapse biopsies. Intermediate results from the larger cohort by Kridel *et al* (n= 118) permitted the authors to consider *EZH2*^{Y646} as alternatively an ancestral or a descendant event. Limited data are available about intratumor/spatial heterogeneity of the mutation but 3/4 mutations retrieved in the Araf *et al* study were actually shared between biopsies from different topography [313].

EZH2^{Y646} was initially reported to have no specific impact, either on OS or on time to transformation in 221 patients [314]. However, data from the PRIMA cohort (155 patients) showed a positive association of both the presence of the mutation and/or gain of *EZH2* found at the time of diagnosis with longer PFS compared with WT cases. Of note, this difference was no

longer significant when assessed after maintenance treatment by Rituximab [302]. Similarly, *EZH2* mutation confers a better prognosis according to m7FLIPI, but one should keep in mind that this score has been established in biopsies taken within one year before R-CHOP treatment [196]. The reason why *EZH2* mutation would be intrinsically associated with less aggressive disease remains unclear. Modified ME observed in mutant cases may still provide a clue (cf. II.3.2.4 and [315]).

II.3.2.3 Impact of the mutation on *EZH2* mRNA & EZH2 protein expression level

No difference was observed as for *EZH2* mRNA expression level between mutant and WT cases, either in FL samples (n mutant = 6, n WT= 10) or in t(14;18) positive *EZH2* mutant GC B cell lymphoma cells [314]. Similarly, EZH2 protein expression level assessed by IHC (using Carlsbad antibody) as well as H3K27me3 levels (using Abcam antibody) remained heterogeneous within samples and couldn't segregate mutant from WT tumors in the same study. To date, no antibody targeting either WT or mutant EZH2 allows a proper assessment of the mutational status in patients samples and the identification of the presence of *EZH2* mutation relies on molecular testing.

However, a more recent publication of a larger cohort showed contrasting results as both mutation and gain of *EZH2* were correlated with a significant increase in transcript level as well as a significant increase of the IHC staining (using CST antibody) in cases with a gain of *EZH2*. Interestingly, a log2 transformed integrative score reflecting H3K27me3/ H3K27me2 ratio was actually higher in mutant cases and displayed a high sensitivity to predict the mutational status of *EZH2* [302].

II.3.2.4 Functional insights of EZH2^{Y646} in follicular lymphoma biology

From the first publication reporting *EZH2* mutation, its putative oncogenic role has been suspected of relating specifically to B cell biology, in opposition with more general mechanisms reported in EZH2 overexpressing solid tumors [242].

RNA-Seq data subsequently provided a hint as to the biological paths altered upon *EZH2*^{Y6Y6}: Initial transcriptomic findings in FL showed a weak enriched signature of 106 differentially expressed genes between 18 mutant and 51 WT cases, with high tumor content and VAF for *EZH2* mutation >17% [301]. Huet *et al* further reported around 1200 activated and 1005 repressed genes from a larger series (PRIMA cohort). In spite of a 20-fold difference, transcriptomic signature in both studies tended to overlap significantly. Interestingly, the latter

publication also reported a major overlap between transcriptomic signatures of *EZH2* mutant and *EZH2* gained cases. Differentially repressed genes in both publications encompassed MYC target genes as well as cell cycle checkpoints [302].

Indeed, abundant literature further unveiled the critical role for *EZH2*^{Y646} in cell cycle control deregulation. Long before the initial discovery of the mutation, one publication reported increased EZH2 expression, along with BIM1, as a marker of proliferation in lymphomas [98]. Ten years later, the first integrative study assessing EZH2 binding and H3K27me3 deposition confirmed a common repressive signature between human centroblasts and hESCs, providing evidence for *EZH2* to maintain dedifferentiation. In lymphoma cells SUDHL4, the cell cycle inhibitor *CDKN1A*, a specific target of *EZH2*, was upregulated then inducing cell cycle arrest upon EZH2 siRNA mediated silencing. It is worth mentioning here that at the time of publication (2010), the authors didn't underline the fact that SUDHL4 actually carries *EZH2*^{Y646S} [316].

This mechanism was further confirmed in light of the knowledge about the mutation by prodigal works from Béguelin *et al* who described in depth the implication of *Ezh2* in inducing GC hyperplasia in Cγ-Cre mice. Indeed, GC hyperplasia is believed to recapitulate a premalignant sate for FL in mouse models. *Ezh2*^{Y641F} conferred a growth advantage and induced tumor progression in a Bcl2+ background through repression of *Cdkn1a.* Moreover, *Ezh2*^{Y641F} allowed the formation of de novo "poised" chromatin domains (i.e. genes harboring a combination of both active and repressive histone marks allowing rapid induction after removal of the repressive marks), especially at genes involved in exit of the GC [296].

Authors from the same group further refined their observations demonstrating that *CDKN1A* repression could also be mediated *via* recruitment by *EZH2*^{Y646} and Bcl-6 of a non-canonical PRC1/BCOR complex, containing CBX8, thus underpinning the potential interest in targeting both Bcl6 and EZH2 in FL [259]. Finally, combining experiments conducted on 3D germinal center organoids and same Cγ-Cre mice models allowed the description of a regulatory loop including phosphorylation of Rb and release of E2F1 to eventually repress CDKN1A and permit un-impaired cell cycling, in turn allowing the GC B cell to undergo the cycle burst "freely" [297].

Another significant step was made with the thorough mechanistic studies by Souroullas *et al* performed in both isogenic melanoma and lymphoma mice models, where *Ezh2*^{Y641} was knocked in. As previously reported, *Ezh2*^{Y641} promoted lymphoma onset in Bcl2 rearranged B cells. RNA-Seq in mutant cells showed a bimodal signature with an equivalent number of up and downregulated genes compared to WT. Repressed genes were at least partially identified as being the targets of PRC2 and encompassed, among other genes targeted by Myc, suggesting the putative oncogenic activity of *Ezh2*^{Y641} as an amplifier of *Myc* effect in B cell lymphomagenesis [294]. These results were actually in line with previous data that demonstrated

the cooperation between Myc and Ezh2 Y641 to promote lymphomagenesis in Eμ-Myc transgenic mice [317]. Of note, full abrogation of PRC2 in these mice accelerated lymphomagenesis [318]. Lines of evidence are actually converging towards a striking oncogenic interplay between *EZH2* and *MYC*: MYC is transiently expressed at the beginning of the GC reaction. Moreover, MYC is silenced by Bcl6 in proliferating centroblasts. Therefore, the combination of activated *MYC* driven signaling pathways in ever cycling *EZH2*^{Y646} cells might efficiently promote tumor growth [319–321].

The most compelling mechanistic finding from the Souroullas *et al* study was the identification of an unexpected deposition pattern of H3K27me3 in the mutant cells, that appears redistributed along the genome instead of increased at pre-existing peaks as it could have been expected: indeed, a majority of highly and focal H3K27me3 peaks in WT cells were actually lost in isogenic mutant B cells and changed into broader and lower peaks (***Figure 15***). Restoration of high peaks by the EZH2 inhibitor gave credit to the concept of a neomorphic rather than a hypermorphic enzyme. This “spreading” of H3K27me3 led to an increase in gene bodies at the whole genome level and thus explained the decreased expression of related transcripts in mutant cells. Upregulated genes tended to be depleted in H3K27me3 around the newly enriched promoters [294]. Of note, Berg *et al* had reached similar conclusions in *Ezh2*^{Y641} transgenic Eμ-Myc mice focusing on a subset of genes [317]. It was later hypothesized that this preferential spreading around the promoters could rely on their proximity with topologically associated domains [322].

Recently, Béguelin *et al* elegantly bridged mechanistic and functional assays in conditional Cγ-Cre mice models, showing not only that *Ezh2*^{Y641} led to an expansion of the centrocyte population in the GC along with a diminution of the centroblastic population, reflecting the trapping of mutant cells in the light zone of the GC, but also the skew in the ME crosstalk of mutant B cells with a decreased reliance on Tfh cells interaction, in parallel with a switch to preferential interplay with FDC. Mutant FL cases had an intact and expansive FDC network (usually a hallmark of low-grade FL) unlike the *EZH2* WT cases where it appeared disrupted, regardless of the grade in this series (***Figure 15***).

Overall, those changes related to repressed gene expression by increased H3K27me3 and were in favor of a bigger picture where *Ezh2*^{Y641} generates a permissive immune niche [315]. Of note, no correlation could be made between upregulated genes and H3K27me3 deposition pattern in this study.

Interestingly, *EZH2*^{Y646} DLBCL (which to some extent may be comparable with FL biology as for the role of *EZH2*) displayed a reduction of MHC class II molecules, contributing in turn to immune escape, echoing the role of *CREBBP* mutations aforementioned [323] (***Figure 15***).

In total, it appears clear that not only mutant EZH2 acts through immune modulation of close environment of tumor cells, but also that combinatorial therapies, including both EZH2 inhibition and immunotherapies, are to be further exploited.

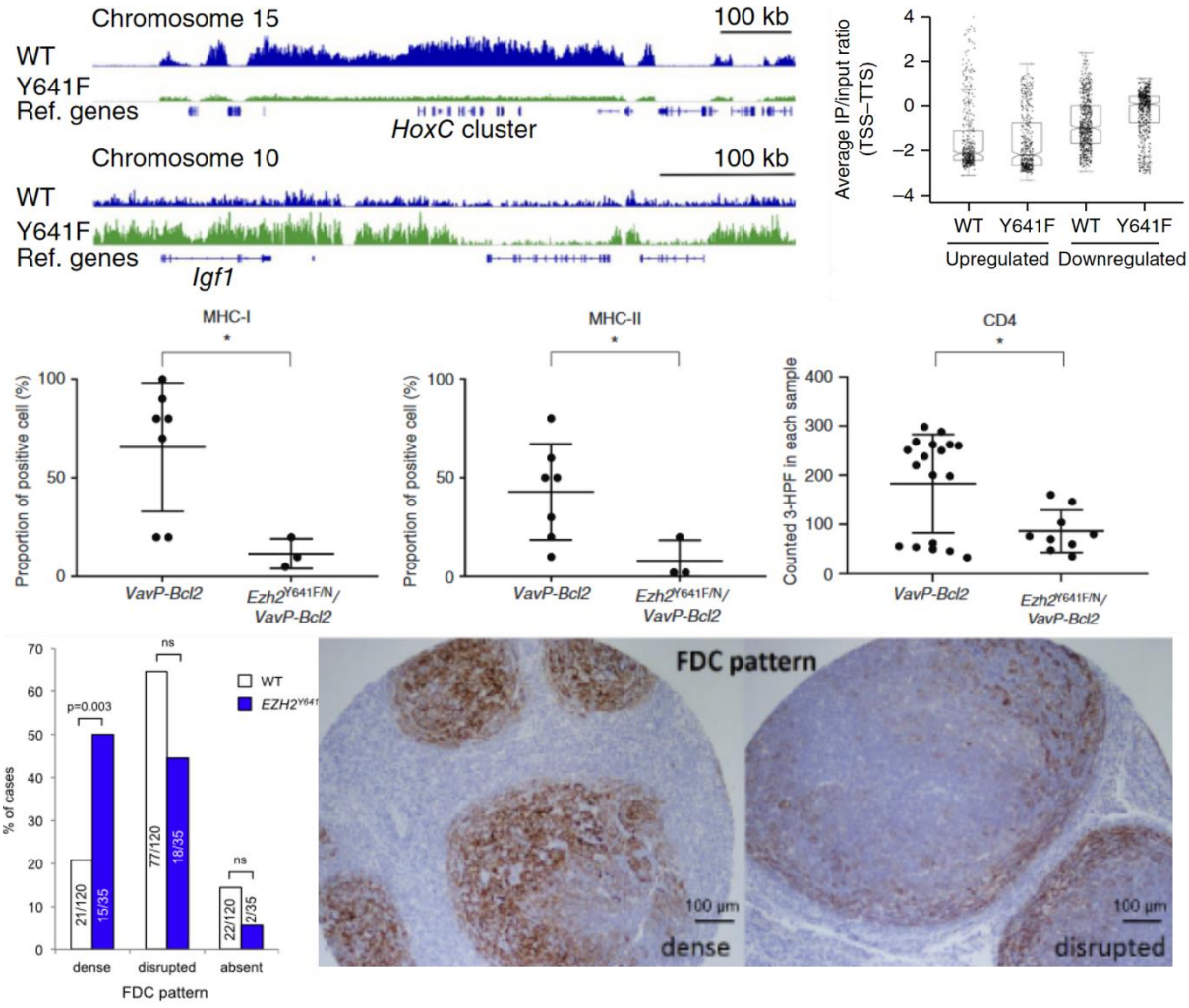


Fig 15: Effects of mutant EZH2 in mice models, and human follicular lymphoma

Redistribution of H3K27me3 upon the mutant background with specific enrichment at downregulated loci, adapted from Souroullas et al (2016) (upper); marked depletion in MHCI/II expressing cells along with CD4 cells in VavP-Bcl2-Ezh2^{Y641} mice, adapted from Ennishi et al (2019) (middle); disruption of the FDC network in EZH2 WT vs mutant FL, adapted from Béguelin et al (2020) (lower)

II.3.2.5 EZH2^{Y646} and DNA methylation

Focusing on FL methylome in a cohort of 164 cases, O'Riain *et al* showed that methylation profile could segregate FL from normal control B cells on the one hand and that, on the other hand, it remains conserved through disease transformation. Interestingly, PRC2 targets in

stem cells appeared to be hypermethylated in FL, mirroring some of the mechanisms of gene silencing in *Ezh2* mutant lymphoma, with no investigation of the effect of the mutation *per se* at the time [324]. However, Bödör *et al* found no major difference of methylation level among 1264 CpG loci between 26 mutant and 220 *EZH2* WT cases, suggesting no specific effect of *EZH2*^{Y646} on DNMT recruitment [314].

II.3.2.6 EZH2^{Y646} and cutaneous melanoma

Out of 929 cutaneous melanoma specimens, Tiffen *et al* reported, in 2016, up to 4.1% of *EZH2* mutant cases, with mutations affected SET domain found in up to 30.9% of the reported mutations. In total, a little less than 1% of the tested cases harbored a mutation of *EZH2* localized in the SET domain (mostly isoform NM_001203247, harboring Y641 mutation) [325]. Lymphoma and melanoma are the only tumor types that have been described as harboring the mutation (parathyroid adenoma display EZH1 SET domain mutation rarely (***Figure 5***) [326]). In addition to somatic mutations, about 5% of the melanoma cases had an amplification in *EZH2* locus, which resulted in an increase of *EZH2* mRNA expression level.

Another difference between *EZH2* mutant melanoma and lymphoma was the correlation between *EZH2* activation (either because of a mutation or an amplification) in melanoma and decreased survival rates in patients.

At the transcriptional level, *EZH2* “hyperactive” melanoma (i.e. mutant and/or amplified and/or with transcript increase) displayed differentially expressed genes compared to non-altered *EZH2* cases, with 55% of repressed genes. Interestingly, 74% of *EZH2* “hyperactive” cases were characterized by differentially hypermethylated loci, which is another striking difference with mutant FL cases. Indeed, further functional assays performed on cells using EZH2 inhibitor GSK126 led the authors to actually conclude that repression of PRC2 targets in mutant melanoma related to both EZH2 hyperactivity as well as hypermethylation, after ruling out the potential implication of altered DNMT activity.

PRC2 target genes highlighted in this study trigger, among others pathways, cell division or stemness maintenance. It is striking to notice that both *CDKN1A* as well as genes involved in antigenic presentation (MHC class II genes) were upregulated upon EZH2 inhibition. A similar observation about *CDKN1A* had previously been reported in melanoma cells [327], echoing observations in mutant *EZH2* GC B cell derived lymphoma (cf. paragraph II.3.2.4). Moreover, from a mechanistic point of view, experiments conducted by Souroullas *et al* in isogenic melanoma and lymphoma cells, and mice in parallel, showed that eroded H3K27me3

deposition pattern as well as the non-univocal transcriptomic consequences in a mutant background were similar in both pathologies [294].

We may hypothesize that EZH2 Y646 leads to a redistribution of H3K27me3 targeting a new set of genes that are shared among cells with various identities. The selectivity of *EZH2*^{Y646} mutation for lymphoma and, to a lesser extent melanoma, highlight the crucial impact of the cellular context and possibly that the aberrantly targeted gene are only oncogenic in those two specific cell types. Besides these observations, a reported case of a patient with metastatic *EZH2*^{Y646} cutaneous melanoma concomitant of a *EZH2* WT FL illustrated that independent oncogenic pathways could cohabit within one individual in regard to *EZH2* mutational status [328].

II.4 Targeting EZH2 in follicular lymphoma

Given the clear association between mutations found in CMG and FL tumorigenesis, investigation of pharmaceutical targeting of these alterations became an obvious direction to take. However, the most common CMG mutations (i.e., *KMT2D* and *CREBBP*) are loss-of-function events and, thus, hardly actionable. As stated by Michael Green, this has made targeting activating *EZH2* mutations the “lowest hanging fruit” [329]. In this section I will review the major steps of the development of EZH2 inhibitors - firstly in their preclinical and subsequently, in their clinical settings.

II.4.1 Preclinical development

3-deazaneplanocin A (DZNep) was the first compound exaggeratedly qualified as an EZH2 inhibitor and used in experimental assays. Indeed, DZNep is a S-adenosyl-L-homocysteine hydrolase (SAH) competitor which leads to an increase of intracellular concentration of SAH, thus limiting all methyltransferases reactions [330]. This compound is however not specific to EZH2. Moreover, despite a proven efficiency in various *in vitro* and *in vivo* models potentially related to PRC2 inhibition, DZNep has a very short half-life and its toxicity profile in animals limited its broad usage [331,332]. UNC1999 was the first orally bioavailable EZH2 inhibitor, effective on both mutant and WT enzyme, yet with a 10-fold less potency for EZH1 inhibition [333].

A breakthrough came with the advent of SAM-competitive small molecules selected upon high-throughput pharmacologic screens: these compounds share as a common mechanism the binding to the SAM pocket of *EZH2* SET domain, with mutual exclusivity with SAH.

Knutson *et al*, from Epizyme company, announced in late 2011 the development of EPZ005687, a pyridinone containing chemotype that showed >500 fold selectivity for PRC2 inhibition over 15 other protein methyltransferases with a selectivity for EH2 over EZH1-containing PRC2 of around 50 fold [334]. EPZ005687 similarly inhibits both WT and mutant EZH2 in lymphoma cells in a dose-dependent manner without disrupting the protein-protein interactions within PRC2 complex. Of interest, EPZ005687 was both cytostatic and cytotoxic in EZH2 mutant cells while, in EZHT WT cells, only cell cycle arrest was reported, suggesting that EZH2 enzymatic activity is a genetic driver in the mutant but not WT condition. Additionally, expression array analysis showed that EPZ005687 treatment induced de-repression of PRC2 targets involved in differentiation.

In parallel, McCabe et *al* developed GSK126, another small molecule wherein biochemical and pharmaceutical properties were globally similar to EPZ005687 [335]. The McCabe study, however, added a layer of information with the realization of H3K27me3 ChIP-seq on lymphoma cells, showing an enrichment of the mark at the responsive genes (i.e. genes upregulated upon inhibition), thus providing a mechanistic insight into the pharmacologic impact of the drug.

A third highly selective SAM competitive inhibitor, EI1 announced early in 2012 by Qi *et al*, inhibited both WT and mutant EZH2 in lymphoma cells [336]. Similarity as observed with the previously mentioned compounds, EI1 caused cytostatic and cytotoxic effect in mutant *EZH2* cells only.

Transcriptomic analyses performed in lymphoma cells with all the 3 molecules (EPZ, GSK, EI1) revealed both up and down regulation of genes upon inhibition: responsive genes as expected, happened to be more numerous than down regulated genes and mostly consisted in PRC2 targets. However, in the McCabe study, the overlap of de-repressed genes within 6 mutant cell lines was very limited (35 genes) showing the importance of the genetic context of each individual cell line before an ontological mechanism strongly shared to mutant cells.

Downregulation of genes upon inhibition may appear as a counterintuitive consequence given the repressive nature of *EZH2*, but a couple of observations argue against a direct mechanism of action of the inhibitor: Indeed, in the McCabe studies, the cell lines that showed substantial gene downregulation were those with the highest number of responsive genes. Moreover, time course RNA-Seq assays performed in EI1-treated cells showed that gene repression happened significantly later after gene upregulation. Together, both these observations

favored the hypothesis for the downregulated genes to be secondary targets of responsive genes, rather than a direct consequence of PRC2 inhibition. However, repressed genes may be considered at the functional level in the assessment of the response to EZH2 inhibitors in patients, as resistance may stem from tumor suppressor gene inactivation.

EPZ005687 was further improved and became EPZ6438, subsequently named Tazemetostat, which is a more potent EZH2 inhibitor, with better pharmacokinetics properties (i.e. good oral bio availability) than EPZ005687. Tumor growth reduction was observed in mice carrying SMARCB1-mutant rhabdoid tumors treated with Tazemetostat, along with a decrease in H3K27me3. Tazemetostat induced not only apoptosis, but also favored differentiation, indicating a release of differentiation blockade as a mechanism of action in tumor cells upon treatment, in line with the concept of synthetic lethality [337]. Similar observations were further made in *EZH2* mutant lymphoma PDX [338]. Interestingly, RNA-Seq performed on treated lymphoma cells further highlighted a potentially targetable dependency on BCR activation [339].

Indole-based EZH2 inhibitors (CPI169 and its trifluororethylpiperidine analog CPI1205) were developed by Constellation company. *In vivo* studies showed antitumor activity, yet selectivity for EZH2 over EZH1 remained limited [340–342]. Finally, Pfizer lactam-derived EZH2 inhibitor PF06821497 showed significant anti-tumor activity in *EZH2* mutant PDX, but to date, literature about this compound is parsimonious [343].

II.4.2 Clinical development

Promising preclinical results for Tazemetostat allowed the realization of a phase 1/2 first in human study (NCT01897571) enrolling a total of 64 patients (21 with B cell lymphoma, including 13 DLBCL (2 *EZH2* mutant) and 7 FL (no *EZH2* mutant cases), along with solid tumors, SWI/SNF sub-unit *INI1* or for SMARCA4 deficient) [344]. Tazemetostat showed favorable pharmacokinetic and pharmacodynamic profiles as well as a tolerance profile; and the dose of 800 mg twice daily was recommended for further clinical investigations. The direct effect of EZH2 inhibition was further appreciated through the reduction of H3K27me3 amount assessed by immunohistochemistry in skin punch biopsies, realized after 28 days of treatment. Additionally, RNA-Seq performed in a biopsy after treatment from one patient with *INI1*-deficient tumor showed a decrease in *EZH2* transcript expression, as well as a differential expression of *EZH2* target genes, when compared with pretreatment expression levels. But available data remains limited to infer the transcriptional consequences of EZH2 inhibition in human tissues. Out of the 21 patients included with lymphoma, 8 underwent an objective response, including 3 complete responses. Of note, the only patient with *EZH2* mutation (in a

DLBCL) experienced a partial response for at least 16 months. Although promising, these results, drawn from a limited number of patients, subsequently required further validation for FL patients, specifically in regard to the *EZH2* mutational status.

The phase 2 trial, ensuing from the first part of NCT01897571, encompassed 5 different arms, including one with 99 patients with refractory/relapsing FL: 45 mutant and 55 WT patients [345]. Results from the study were non-ambiguous: the objective response rate in *EZH2* mutant FL reached 69% and, surprisingly, 35%, in WT patients (***Figure 16***). PFS was slightly longer in the mutant group, and complete response was observed in 13% of the patients from this group vs. 4 % in the WT group. Severe adverse events were merely anecdotal, and tolerance was very good compared to other compounds licensed for 3rd. line therapy - such as, for example, PI3K inhibitors or Lenalidomide/Rituximab. Interestingly, performance of Tazemetostat normalized on POD24 as a marker of poor outcome showed equally favorable results in patients with a good or bad prognosis (63 and 25% of ORR in POD24 mutant and WT patients respectively).

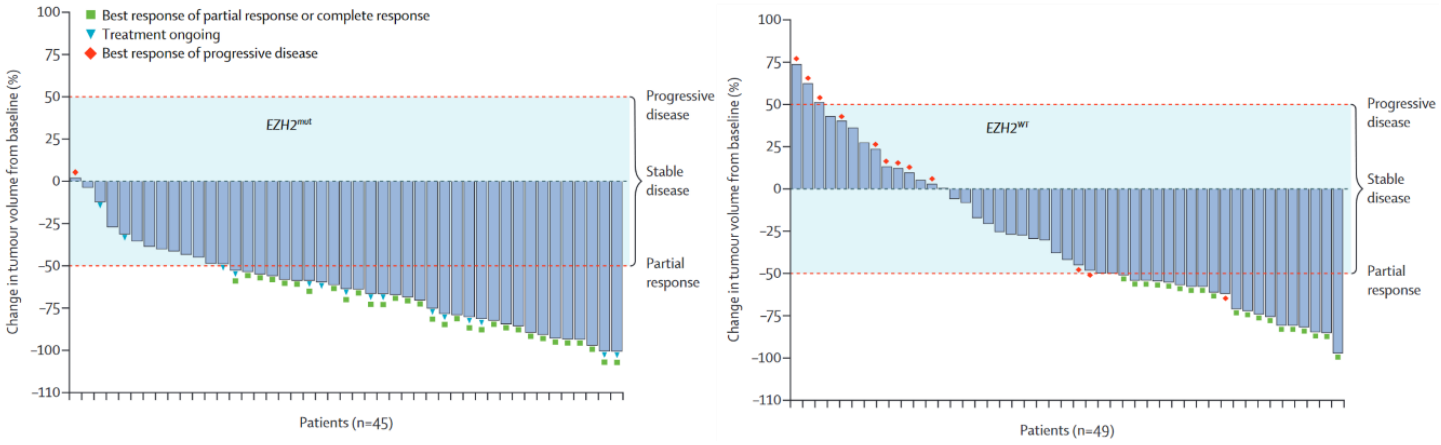


Fig 16: Response to Tazemetostat in *EZH2* mutant and WT FL patients, results from a phase 2 trial, *from Morschhauser et al (2020)*[345]

In June 2020, Food and Drug Administration granted accelerated approval to Tazemetostat (TAZVERIK, Epizyme, Inc.) for treatment of adult patients with relapsed or refractory FL, after 2 lines of prior systemic therapies and whose tumors were positive for an *EZH2* mutation, as detected by Cobas *EZH2* Mutation Test (Roche Molecular Systems, Inc.) as a diagnostic companion test. To date, no evidence supports discontinuing treatment by Tazemetostat, if

well tolerated. According to Epizyme, out of 729 patients treated with Tazemetostat, 0.7% developed myelodysplasia or acute myeloid leukemia.

The association of Tazemetostat with R-CHOP presents safety and pharmacological profiles comparable with administration of uniquely R-CHOP in 60-80 years old patients with newly diagnosed DLBCL [346]. Preliminary efficacy data is encouraging, but further investigations in a phase 2 study are warranted. Indeed, interim results available so far in DLBCL patients treated with Tazemetostat are not as obvious as in FL, considering *EZH2* mutational status.

Attempts have been made to predict responses to Tazemetostat beyond the sole presence or absence of *EZH2* mutation. To date, it appears that activating mutations in *MYD 88* in *EZH2* WT patients was associated with response, whereas patients with mutation in *HIST1H1E* or *MYC* showed a trend for non-response. In addition, a combinatorial pattern consisting of WT *MYC* and/or *HIST1H* WT, with a mutation in *STAT6* or *MYD288* , showed a better response [347]. This algorithm remains to be validated.

With preclinical observations, notably in ovarian tumor cells, showing a synergistic effect of PDL1 inhibitors, in addition to EZH2 inhibitors following an enhanced interplay between tumor cells and surrounding T cells upon EZH2 inhibition, combinatorial studies with Tazemetostat and Atezolizumab were initiated, and then stopped in DLBCL. Indeed, inhibiting EZH2 may have alternate tumor cell–independent therapeutic activity, but the complexity of *EZH2* related cross talk and the imperfectly established role of nonmalignant immune cells, may make it difficult to predict the efficacy of EZH2 inhibition in FL, based on mutational status only.

As of August 2021, 30 clinical trials are registered with the aim to assess Tazemetostat, either alone or in combination with various compounds, such as Bcl2 inhibitor Venetoclax, or BCR modulators, in FL or other tumor types.

EZH2 inhibitors, other than Tazemetostat, have been assessed in clinical cohorts as well, but so far with less success: GSK126 was assessed in non-Hodgkin lymphoma, multiple myeloma and solid tumors [348]. The absence of significant efficiency did not justify further investigation in lymphoma, according to *EZH2* mutational status, as was originally planned. GSK126 in combination with pomalidomide showed enhanced anti-tumor efficiency in GCB DLBCL cells, which may open the door for upcoming clinical investigations [349].

CPI-1205 is well tolerated and has an acceptable toxicity profile, with evidence of anti-tumor activity and access to the target. Thus, an expansion phase in GCB-DLBCL, or FL without *EZH2* mutation, was initiated with CPI-1205 at 800 mg BID; and subsequently TID, as well as

trials assessing CPI-1205 in combination in solid tumors are ongoing. No results are available to date. One phase 1 trial with PF06821497A is currently enrolling patients with FL among other tumor types.

Some mechanisms of resistance to EZH2 inhibition are proposed: secondary mutations in *EZH2* SET domain have been shown to confer resistance to the EZH2 inhibitor in DLBCL cell lines, which could, in theory, be overcome with EED inhibition [350]. However, given the highly toxic profile of EED inhibitors currently available, this hypothesis remains purely speculative. Upregulation of alternate survival pathways, such as IGF1R, MEK or PI3K pathways, induced repression of apoptotic gene expression *TNFSF10* and *BAD*, thus conferring resistance to EZH2 inhibitor in the same cells. A recent publication showed that Ikaros protein inactivation (*IKZF1* gene, identified *via* a synthetic lethality-based CRISPR/Cas9 screen) sensitizes DLBCL cells to EZH2 inhibitor, regardless of the presence of the mutation [351]. One study underlined the specific effect on chromatin of targeting EZH1 in addition to EZH2, in both *EZH2* WT and mutant DLBCL cells, thus providing a rationale for clinical trials evaluating the EZH1/2 inhibitor in FL [352].

Finally, as co-occurrent *KMT2D* and/or *CREBBP* mutations might impact the H3K27me3 landscape of *EZH2* mutants or WT FL tumors, combined therapies targeting one altered CMG or more thus appears a legitimate path to follow in the field of pharmaceutical compound development. As an example, pan HDAC inhibitors have been reported with high a response rate in FL and could eligibly target cases with *CREBBP* inactivating mutations [221]. Moreover, HDAC3 inhibitor has been shown to induce re-activation of genes silenced by *Crebbp* loss, including MHC class II genes [353]. Given the similitude observed with EZH2 inhibition in *EZH2* mutant cells, it would be fair to hypothesize a synergistic action of using these compounds in co-mutated cases. However, there is an absence of both preclinical models and translational studies to address the question of the role of multiple mutations in CMG in FL biology.

OBJECTIVES

Genetic mutations affecting histone modifier Polycomb Repressive Complex 2 (PRC2) are described in a wide range of tumor types. The critical influence of altered PRC2 in driving tumorigenic processes as well as its surrogate role as a therapeutic target have raised major interest over the past two decades and led to remarkable progress at the clinical level. Indeed, patients with mutant *EZH2* follicular lymphoma (FL) are now eligible for a targeted therapy in regard of their mutational status. However, in spite of the significant amount of discoveries that have been made about each of its alterations encountered in a specific context, several aspects of the unique biology of PRC2 remain uncovered. During my PhD, I sought to tackle two unanswered questions about its multifaceted role in cancer.

First, given the high cell type specificity of each individual lesion affecting PRC2 in cancer, little is known about the precise mechanisms and consequences of these mutations independently from the genomic and chromatin environment in which they arise. We thus hypothesized that modeling the most frequent of these alterations in one single isogenic cell line would help us to better understand their respective impact on chromatin landscape and transcriptome. In the first part of the following section (Results, paragraphs 1, 2 and 3), we took advantage of engineered cells recapitulating PRC2 alterations found in cancer (*Eed*-KO, *Ezh2*-KO, H3.3K27M, *Ezh2*^{Y641F}) using integrative RNA-Seq/CUT&RUN-Seq as well as HMT assays. We compared, side by side, the relative degree of impairment of PRC2 in mutations inducing a loss of its enzymatic activity. We then sought to refine the knowledge about the dual effect of *Ezh2*^{Y641} in reprogramming its chromatin environment. Finally, we dissected the mechanisms underlying the differential response to EZH2 inhibitor observed between mutant and wild type cells, mirroring what is observed in patients with FL.

Second, investigations about the biology of mutant chromatin modifiers in FL, and more specifically *EZH2*^{Y646}, are numerous in preclinical models but translational studies on this topic mostly rely on genomic and transcriptomic data. We reasoned that studying chromatin changes in tumors samples was crucial in a pathology such as FL. We then intended to build up a collection of longitudinal clinical samples from patients with both *EZH2* mutant and WT FL, along with clinical history. In the second part of the following manuscript (Results, paragraphs 4 and 5), we introduce a unique cohort of 32 cases for which we collected multi-omic data (genomic and CNV data, RNA-Seq, H3K27me3 ChIP-Seq). Ultimately, we performed a coupled ChIP-Seq/RNA-seq analysis to objectify the consequences of the mutation on H3K27me3 deposition in FL samples.

In summary, the aims of this PhD project were to broaden the understanding of PRC2 alterations in cancer in general, and of *EZH2*^{Y646} mutation in follicular lymphoma in particular. In order to do so, we undertook a double approach, mechanistic on the one hand, translational on the other hand. In the next section, I will present our results combined in one manuscript for an article that is still in preparation to date.

In addition to this project, I had the chance to participate in a publication focusing on the relative contribution of canonical and non-canonical activity of PRC2 to gene silencing and on the proliferation-dependent redundancy that exists between EZH1 and EZH2 in cancer.

Wassef M, Luscan A, Aflaki S, Zielinski D, Jansen PWTC, Baymaz HI, et al. **EZH1/2 function mostly within canonical PRC2 and exhibit proliferation-dependent redundancy that shapes mutational signatures in cancer.** Proc Natl Acad Sci U S A. 2019;116(13):6075‑80.

Moreover, I was able to terminate a publication with my co-director, Dr Vincent-Salomon, that is beyond the scope of PRC2. This article is a comprehensive study about the specificity of medullary breast carcinoma in regard of other basal-like breast cancers.

Romero P, Benhamo V, Deniziaut G, Fuhrmann L, Berger F, Manié E, et al. **Medullary breast carcinoma, a triple-negative breast cancer associated with BCLG overexpression.** Am J Pathol. 2018;188(10):2378-91

These two articles are added in the appendix of this thesis manuscript.

RESULTS

Title: Modeling PRC2 alterations in cancer, a window toward a better comprehension of *EZH2* mutation in follicular lymphoma *(Manuscript in preparation)*

Authors : Pierre Romero[1], Laia Richart[1], Setareh Aflaki[1], Megan Burton[1], Audrey Michaud[1], Julien Masliah-Planchon[2], Frédérique Kuhnowski[3], Céline Méaudre[4], Michel Wasssef[1], Anne Vincent-Salomon[4], Raphaël Margueron[1]

Affiliation :

[1] Institut Curie, CNRS UMR3215, INSERM U934, Paris Sciences et Lettres Research University, Sorbonne University, 75005 Paris, France

[2] Institut Curie, Pharmacogenomics Unit, Department of Genetics, Paris Sciences et Lettres Research University, 75005 Paris, France

[3] Institut Curie, Department of Clinical Hematology, Paris Sciences et Lettres Research University, 75005 Paris, France

[4] Institut Curie, Department of Pathology, Paris Sciences et Lettres Research University, 75005 Paris, France

Abstract

Chromatin plays a determinant role in the regulation of transcriptional programs. Mutations found in genes that modulate its conformation are of prime relevance in the biology of a wide range of cancers. Notably, Polycomb Repressive Complex 2 (PRC2) is a chromatin modifier altered in various tumor types. PRC2 catalyzes the methylation of H3K27, associated with inactive genes. The consequences of these alterations are, however, highly context-dependent: Indeed, some of them result in abrogation of the deposition of H3K27me3, whereas other, namely the mutation Y646 of the catalytic sub-unit of PRC2 *EZH2* and present in up to 25% of follicular lymphoma (FL), leads to both an increase and a redistribution of H3K27me3. The phenotypic changes induced by *EZH2*^{Y646} in FL cells remarkably drive a better response than *EZH2*^{WT} FL to Tazemetostat, an EZH2 inhibitor that has been recently licensed for the treatment of FL. However, the precise mechanisms underlying the consequences of both loss and gain-of-function of PRC2 in cancer remain partially unknown. In order to gain insight on this question at both fundamental and translational levels, we conducted an integrative study relying on two complementary approaches. In a mechanistic approach, we characterized isogenic cell lines recapitulating a collection of PRC2 alterations found in cancer (*Eed*-KO, *Ezh2*-KO, H3.3K27M) and show their unambiguous yet, uneven, degree of PRC2 inhibition. In contrast, *Ezh2*^{Y641F} induces an increase and a redistribution of H3K27me3 that change the transcriptomic response to EZH2 inhibition. Interestingly, both loss and gain-of-function alterations of PRC2 are associated with a global rewiring of the chromatin environment. Further ChIP-Seq/RNA-Seq profiling of a unique collection of *EZH2* mutant and WT FL tumor samples confirms the epitranscriptomic changes induced by the mutation and allows us to objectify an impact of H3K27me3 redistribution over B cell-specific enhancers in *EZH2* mutant cases.

Introduction

Chromatin plays a determinant role in the regulation of transcriptional programs. This nucleoproteic structure is tightly controlled by a variety of enzymes whose function is often altered in cancer [1,2]. Among them, the Polycomb Repressive Complex 2 (PRC2) catalyzes mono, di and tri methylation of the lysine residue 27 of histone H3 (H3K27me1/2/3) through the lysine methyltransferase activity of Enhancer of Zeste Homologue 1/2 (EZH1/2). In addition to its catalytic subunit EZH1/2, the core PRC2 complex also includes Embryonic Ectoderm Development (EED), Suppressor of Zeste 12 (SUZ12) and Retinoblastoma Associated protein 46/48 (RbAp46/48). PRC2 and H3K27me3 are tightly associated with transcriptional silencing. Deregulation of PRC2 in the form of EZH2 overexpression is often detected in solid tumors, particularly in those of poor prognosis, thus a causative link was drawn between PRC2 hyperactivity and tumor aggressiveness [3]. Nonetheless, this link has been weakened by evidence that EZH2 expression is coupled to cell-cycle, therefore its overexpression may reflect a more proliferative or dedifferentiated state with no clear functional impact on cancerous transformation *per se*. Yet, PRC2 activity has also been reported to be modulated by genetic mutations affecting either its different subunits or its H3 substrate in a cancer type-specific manner [4]. Indeed, PRC2 has a multifaceted activity in cancer being either tumor suppressor or oncogene depending on the context.

Both EED and SUZ12 are required for the stability of PRC2 and their loss causes dismantling of the complex and abrogation of its enzymatic activity. In this way, a substantial number of malignant peripheral nerve sheath tumors (MPNSTs), as well as a small proportion of melanoma and of glioblastoma multiform, present deletions for either *EED* or *SUZ12*, suggesting in these cases the tumor suppressive function of PRC2 [5]. Hematological malignancies arising from various developmental stages of blood cells, of both myeloid and lymphoid lineages, also display abolished PRC2 activity [6]. The landscape of PRC2 alterations in this subgroup of tumors is complex as all members of the core complex can be impaired along with variable degrees of transcriptomic and epigenomic consequences. Chromatin repression by PRC2 may also be hampered by mutated histones (a.k.a. “oncohistones”): histone H3 mutation H3K27M is detected in midline glioma, a highly-lethal pediatric brain tumor, and leads to PRC2 partial loss of function [7]. Recently, the H3K27M-mimmicking protein EZHIP has been reported to be abnormally expressed in a majority of posterior fossa group A - ependymoma and results in a loss of PRC2 activity similar to H3K27M itself, without incorporation to chromatin [8]. The exact mechanisms underlying oncogenesis induced by oncohistone and oncohistone-like proteins remain unraveled to date. These examples thus

illustrate the diversity of the channels through which PRC2 loss-of-function can drive tumorigenesis.

Opposite to loss-of-function alterations, the mutation *EZH2*^{Y646} was described a decade ago in germinal center (GC)-derived B cell lymphomas, and more specifically in up to ~25% of follicular lymphoma (FL), a malignancy strikingly affected by mutations in chromatin modifying genes [9–11]. We have now robust evidence that increased activity of PRC2 resulting from *EZH2*^{Y646} leads to a differentiation blockade of B cells along with a reprograming of the microenvironment, both crucial for lymphomagenesis [12–14]. Works on mechanisms of action of *EZH2*^{Y646} have yielded unexpected conclusions: rather than catalyzing increased deposition of H3K27me3 at canonical target loci, the mutant enzyme globally redistributes H3K27me3 along the genome, which results in equivocal transcriptomic consequences [15]. This unconventional change-of-function of PRC2 might impact the effectiveness of epidrugs that inhibit EZH2 activity and thus the therapeutic strategy. In fact, Tazemetostat has shown higher anti-tumor efficacy in FL patients harboring the *EZH2* mutation than in those WT for *EZH2* [16]; the underlying mechanisms remaining opaque

Thus PRC2 is crucial to the control of transcriptomic programs frequently altered in a wide range of tumor types. Each individual lesion of PRC2 has been under the scope of abundant investigations but, to our knowledge, no study has so far attempted to evaluate their precise consequences in one single isogenic model. Here, we report the characterization of isogenic cell lines recapitulating a collection of PRC2 alterations found in cancer: *Eed*-KO, *Ezh2*-KO, H3.3K27M, and *Ezh2*^{Y641F}. H3.3K27M, Ezh2-KO and Eed-KO, despite disrupting PRC2 activity through distinct mechanisms, appear to be mostly distinguished by the degree of PRC2 inhibition. In contrast, *Ezh2*^{Y641F} induces an increase and a redistribution of H3K27me3 that change the transcriptomic response to EZH2 inhibition. Profiling by ChIP-Seq of a unique FL cohort built over a period of 25 years and including both WT and EZH2-mutant samples further confirms the rewiring of the H3K27me3 landscape in the presence of the mutation. In addition to the differential accumulation of H3K27me3 on promoters and gene bodies, we report increased deposition of H3K27me3 over B cell-specific enhancers in EZH2 mutant cases and speculate that this could participate in the transcriptomic deregulation associated to the mutation.

Results

H3.3K27M, Ezh2-KO and Eed-KO affect H3K27me3 deposition to variable degrees

PRC2 activity is impaired in a wide range of human malignancies. Mechanisms resulting in its loss-of-function are heterogeneous and their effect on tumorigenesis has been extensively explored in context-specific studies. However, a direct and comprehensive comparison of their epitranscriptomic consequences in a single cellular model has not been done to our knowledge. Toward this end, we performed a multi-approach characterization of various cancer-associated PRC2 mutations using immortalized mouse embryonic fibroblasts (iMEFs) that are either modified *via* 4-OHT (TAM)-inducible Cre-mediated deletion of *Ezh2* or by CRISPR-Cas9-mediated genome editing (***Fig 1a, S1a***). We previously showed that *Ezh2* deletion has no impact on proliferation in this model [17], thus ruling out cell cycle-related effects. We first analyzed how each of these mutations modulates the global levels of H3K27 methylation (H3K27me) by Western blotting (***Fig 1b***). While deletion of *Eed* fully abrogates H3K27me, as observed in MPNSTs, expression of H3.3K27M has the mildest consequences with a partial reduction of H3K27me2/3 and echoes the behavior of PRC2 in midline gliomas. Deletion of *Ezh2,* found in several hematological cancers, has an intermediate phenotype with a robust reduction of H3K27me3 but less pronounced consequences on H3K27me2. The differences between *Eed* and *Ezh2*-KO cells likely reflect the partial compensation of *Ezh2* deletion by its paralog *Ezh1*.

We then investigated how these alterations of H3K27me deposition translate at the transcriptomic level by performing RNA-Seq. Given the role of PRC2 in transcriptional repression, we focused on the significantly up-regulated genes (FDR < 0.05 and Log2 Fold Change [FC] > 1) (***Fig 1c, S1b***). 1634 genes are upregulated upon deletion of *Eed*, whereas only 959 and 784 are upregulated upon deletion of *Ezh2* or H3.3K27M, respectively. The three sets of differentially expressed genes show a high degree of overlap, as expected from alterations that gradually impair PRC2 function (***Fig 1c, S1c***). Besides, *Ezh2* deletion in cells that already express H3.3K27M has modest transcriptomic consequences (46 upregulated genes, ***Fig S1b***), thus confirming that H3.3K27M impairs most of EZH2 activity.

Profiling of genome-wide H3K27me3 by CUT&RUN-Seq in *Eed* and *Ezh2*-KO cells showed complete and partial abrogation of H3K27me3 deposition, respectively (***Fig S1d-e***). Similarly, we observed a substantial attrition of H3K27me3 in H3.3K27M cells compared to WT cells (***Fig 1d-e***). Previous studies have reported a H3.3K27M-related gain of function of PRC2; namely,

de novo H3K27me3-enriched loci [18]. We thus investigated the reduction of PRC2 enzymatic activity using a peak calling algorithm (MACS2). Consistent with our observations at the whole genome level, the number of peaks is drastically reduced in mutant cells when compared to WT (***Fig 1f***) and, at those peaks that are preserved, we report a strong reduction of intensity of H3K27me3 intensity (***Fig 1g***). An anecdotal number of peaks seemed to differentially gain H3K27me3 compared to WT (36/16400). However, careful examination of these peaks suggests artefactual peak detection due to high background in the corresponding region (***Fig S1f***).

In order to define how PRC2 alterations modulate the chromatin landscape, we performed a comparative chromatin-state discovery and annotation analysis in WT *versus* either H3.3K27M or *Ezh2*-KO cells using ChromHMM [19]. Using a panel of CUT&RUN-Seq experiments for histone marks associated with transcriptional activation (H3K36me3, H3K4me3, and H3K27Ac) and repression (H2AUb, H3K27me2, and H3K27me3), we defined 9 ChromHMM states encompassing areas of active and inactive chromatin. A 10th ChromHMM state (Quies/Low) was characterized by the absence or low levels of the assessed histone modifications and represents the vast majority of chromatin (***Fig 1h, S1g***). H3K27me3 is reduced in both mutant conditions across chromatin states. In contrast, changes in H3K27me2 diverge between cell lines: while reduction of H3K27me2 spans all ChromHMM states in *Ezh2*-KO cells, it is compartment-specific in H3.3K27M cells. We could even detect a small increase in H3K27me2 at E9 (Repressed by Polycomb), suggesting that, in the presence, of H3.3K27M H3K27me2 acquires a distribution that is comparable to that of H3K27me3 in WT cells (**Fig 1h**). Similar observations were previously reported with H3K27M or upon expression of EZHIP [20,21]. Another interesting observation is the tendency of mutant cells to gain histone marks associated with active transcription (H3K4me3, H3K27ac, and, in *Ezh2*-KO cells only, H3K36me3). The global enrichment of H3K27ac in cells expressing H3K27M was reported previously [22] (***Fig 1h, S1h-i***), however we noticed that the E2 stat (Actively transcribed TSS) was excluded from this gain of H3K27ac in both mutant cells. This was further confirmed through H3K27ac peak calling analysis: H3K27ac peaks defined in WT cells show major depletion upon H3.3K27M mutation and *Ezh2*-KO, along with enrichment in flanking regions (***Fig S1j***).

Altogether, we have established a unique system to model cancer-associated mutations disrupting PRC2 activity. We show that H3.3K27M leads to a non-equivocal reduction of PRC2 enzymatic activity with a similar phenotype to the one observed upon deletion of *Ezh2.* Finally, PRC2 loss-of-function are associated to a global increase of active histone marks. Noticeably,

in the case of H3K27ac this increase is distributed across the genome, however, H3K27ac peaks tend to show a reduced enrichment.

EZH2 Y641F promotes a global change of chromatin landscape with dual consequences

Next, we engineered *E*zh2$^{Y641F/WT}$ iMEFs using a strategy that concomitantly eliminates the 5' *Lox*P site of the targeted allele, thus rendering it insensitive to TAM-induced recombination (***Fig 2a, S2a***). *Ezh2*$^{Y641F/WT}$ models *EZH2*^{Y646} found in ~25% of FL. Expression of this mutant form of Ezh2 leads to a clear increase of H3K27me3 global levels, as measured by Western blotting, accompanied by a concordant decrease in H3K27me2 (***Fig2b***). Since, *in vitro*, PRC2-EZH2^{Y641F} is unable to methylate recombinant histones, it was proposed that the mutant complex cooperates with its WT counterpart in order to yield a global increase in H3K27me3 [9,23]. However, whether such cooperation occurs *in vivo* is still unclear. To address this question, we treated our *E*zh2$^{Y641F/WT}$ cells with TAM to delete the remaining floxed WT allele; however, this did not alter the total levels of H3K27me3 (***Fig 2b***). Knocking out *Ezh1* did not have any impact either on the global levels of H3K27me3 (***Fig 2b***). These results strongly suggest that, contrarily to previous assumptions, PRC2-EZH2Y641F can catalyze mono-methylation of H3K27 *in vivo*. Several cofactors interact with PRC2 in order to foster and modulate both its recruitment to chromatin and enzymatic activity [24]. This led us to question the role of PRC2 cofactors in the regulation of the mutant complex. Thus, we performed Histone Methyl-Transferase (HMT) assays *in vitro* with reconstituted WT or *EZH2* ^{Y641F}-carrying PRC2 (PRC2 Y641F) complexes along with PRC2 cofactors AEBP2 and/or JARID2 (***Fig 2c***). This assay reveals that in presence of JARID2 and AEBP2, PRC2-EZH2Y641F is now able to catalyze the methylation on a recombinant substrate, although a bit less efficiently than its WT counterpart.

Using CUT&RUN-Seq, we then investigated the genome-wide consequences of this mutation at the chromatin level. We confirmed the global redistribution of H3K27me3 with an erosion in the height of peaks but a broad spreading of the mark (***Fig 2d)***. Consistently with the Western blot results, enrichment for H3K27me2 is close to background when *Ezh2*^{Y641F} is expressed (**Fig 2b & 2e, S2b**). We further investigated the changes in chromatin composition in *Ezh2*$^{Y641F/WT}$ cells in the context of annotated chromatin states defined above (***Fig 2f, S1g***). Erosion of H3K27me3 peaks in the mutant condition is illustrated by the decrease of signal in ChromHMM state E9 (Repressed by Polycomb) (***Fig 2g***). Yet, in agreement with the observations made by Western blotting, H3K27me3 is increased in all other chromatin states,

with a predilection for E4 and E6. While E4 likely corresponds to weak enhancers, E6 is characterized by the sole presence of H3K27me2 in WT cells (***Fig S1 g),*** and both E4 and E6 are depleted for H3K27me3 (***Fig 2g***). This suggests that, upon expression of *Ezh2*^{Y641F}, H3K27me3 tends to be gained at places that were initially decorated by H3K27me2. Another striking observation is the global increase of H3K27ac in mutant cells, with a marked enrichment at H3K27ac pre-bound loci (***Fig 2f, S1h***). Still, we do not observe newly generated peaks in the mutant background. We hypothesize that the H3K27ac landscape in mutant cells could result from the diminution of H3K27me2 which is not fully compensated by H3K27me3, thus leaving many lysine residues (H3K27) unmodified and amenable to acetylation.

The profound reshuffling of H3K27me3 deposition upon expression of *Ezh2*$^{Y641F/WT}$ led us to question whether EZH2^{Y641F} is merely a hyperactive enzyme or else it has *de novo* targeting properties. We reasoned that partial inhibition of PRC2 enzymatic activity should restore a WT-like pattern of H3K27me3 deposition in the first case but not in the second. Of note, a previous report had suggested that the mutant-related pattern was reversible upon partial EZH2 inhibition in melanoma cells, the authors describing restoration of the heights of H3K27me3 domains [15]. We therefore treated the cells with increasing but relatively low doses of the EZH2 inhibitor (EZH2i) UNC 1999 and performed H3K27me3 CUT&RUN-Seq analyses. In order to simplify the data interpretation, the experiment was done on cells expressing only the mutant enzyme (TAM-treated *Ezh2*$^{Y641F/-}$ *Ezh1*-KO cells) *versus* WT control (***Fig 2i***). In the WT setting, we observe a clear reduction of both the height and number of H3K27me3 peaks (***Fig 2i, left & right panels***). Interestingly, the remaining peaks overlap with the ones observed in the H3.3K27M cells and with the previously described nucleation sites, although a proper quantification is made difficult by the different technologies used for both studies (***Fig S2c***) [25]. Mutant cells offer partial resistance to EZH2 inhibition and retain H3K27me3 at higher levels than WT cells (***Fig 2i, left & middle panels***). Still, we observe a global reduction of the broad and eroded peaks, along with a progressive reshaping of the peaks towards a more WT-like pattern, with no concomitant restoration of the heights of H3K27me3 peaks (***Fig 2i, left panel***). This is associated to an increase in the number of H3K27me3 peaks along with EZH2 inhibition, probably reflecting the better segregation between peaks and background (***Fig 2i, right panel***). The discrepancy regarding peak height upon EZH2 inhibition between our results and a previous report might stem from the use of different normalization methods [15]. Indeed, removing the spike-in normalization from our analysis pipeline yields very different H3K27me3 profiles (***Fig S2d***). Nonetheless, partial inhibition of EZH2^{Y641F} does not concentrate H3K27me3 deposition around putative nucleation sites as observed in the WT, instead it homogeneously lowers the enrichment.

Hence, our results suggest that the Y641F mutation of *Ezh2* not only affects the catalytic activity of PRC2 but also the distribution of other histone marks over the chromatin. The sites where PRC2 preferentially catalyzes H3K27me3 are not anymore restricted to the dense CpG islands, but instead encompass broader genomic regions normally decorated by H3K27me2. Even though the mutant enzyme is more active, H3K27me3 is depleted at PRC2 canonical peaks.

Distinct transcriptomic response to EZH2 inhibition of cells expressing Ezh2^{Y641F}

Patients with FL mutant for *EZH2* respond better to therapies based on EZH2 inhibitors than patients with the WT enzyme [16]. This could suggest that EZH2^{Y641F} is more efficiently inhibited by EZH2 inhibitors or, alternatively, that the consequences of the inhibition are distinct in an EZH2 WT *versus* mutant context, the latter inhibiting more efficiently tumor growth. In order to provide a mechanistic rationale for this uneven response to EZH2 inhibitors, we investigated how *Ezh2*^{Y641F} modifies the transcriptional response to EZH2 inhibition. Differential expression analysis of *Ezh2*$^{Y641F/WT}$ cells revealed a similar number of up- and down-regulated genes (993 and 1078, respectively) (***Fig 3a***). Interestingly, Gene Ontology (GO) analysis indicates that both set of genes are enriched for terms that define classical PRC2 targets such as developmental or differentiation-related pathways (***Fig 3b***). This is consistent with the dual effect of *EZH2* ^{Y641F} on chromatin: the spreading of H3K27me3 with a concomitant erosion of its canonical peaks.

Next, we investigated the transcriptomic response to partial or complete inhibition of PRC2 by treating the cells with two doses of UNC 1999: 0.25µM and 2µM. The RNA-Seq analysis indicated that the number of responsive genes (*i.e.* genes upregulated upon EZH2 inhibition) is similar between WT and *Ezh2*$^{Y641F/WT}$ cells for a given dose of UNC1999, but that the nature of these genes is different from one genotype to another (***Fig3c, S3a***). Moreover, Principal Component Analysis (PCA) revealed that most of the inter-sample variability is explained by the genotype (i.e. *Ezh2* status), and that EZH2 inhibition does not reduce the distance between WT and mutant cells as hypothesized (***Fig 3d***). To further illustrate the transcriptional response to EZH2 inhibition, we generated a heatmap of the responsive genes in either WT or *Ezh2*$^{Y641F/WT}$ cells treated with UNC 1999 (***Fig 3e***). A minority of genes are responsive in both cell types ('Common'); unsupervised clustering of the rest reveals 4 main clusters of genes defined by their genotype-specific responsiveness. Intersection of the transcriptional and CUT&RUN-Seq data revealed that, in WT cells, enrichment for H3K27me3 is higher among genes that selectively respond to EZH2i in these cells (clusters 1 & 2 *vs* clusters 3 & 4, ***Fig 3f***).

The reverse is less clear which is consistent with the broad and probably less specific distribution of the mark in the mutant condition.

GO analysis of responsive genes unveiled that, in WT cells, the classical terms associated to PRC2 targets are significantly up-regulated in response to EZH2i (***Fig 3*g, left panel, and S3b-d**). In marked contrast, genes upregulated in mutant cells show limited enrichment for GO terms. Remarkably, there are 5 terms significantly over-represented among genes up-regulated in mutant cells related to antigen presentation (***Fig 3g, right panel and S3b-d***). When exploring the detailed expression data of the genes included in these pathways, it was striking to notice that mice counterpart of human HLA related genes – a majority of the H2 complex gene - were the most differentially activated genes in the mutant cells (***Fig 3h***). Despite the fact that PRC2 has been suggested to silence antigen presentation pathway in order to promote immune evasion of B cell lymphoma cells, making similar observations in a model so distant from B cells was unexpected [26].

We conclude from these experiments that *Ezh2* activating mutation not only modifies the chromatin landscape and the transcriptome of mutant cells but also the response to EZH2 inhibition. Interestingly, we observed upregulation of antigen presentation-related genes upon EZH2 inhibition in mutant cells. Given the relevance of the antigen presentation pathway in the context of tumorigenesis, further epitranscriptomic research is required in a cellular model that more closely recapitulates B cell biology.

A collection of clinical tumor samples enables a longitudinal multi-omic characterization of *EZH2*^{Y646} and *EZH2*^{WT} follicular lymphoma

Our observations on the epitranscriptomic changes induced by *Ezh2*^{Y641} in iMEFs prompted us to apprehend the role of *EZH2*^{Y646} in a human tumor dataset. *EZH2*^{Y646} occurs in combination with other mutations affecting chromatin-modifying genes which, along with the critical influence of the microenvironment, most likely impact its contribution to lymphomagenesis. The data available to date on the role of *EZH2*^{Y646} in FL consists of chromatin immunoprecipitation experiments and functional assays performed in *EZH2* mutated DLBCL cell lines and mouse models. However, no translational study compiling genomic, epigenomic and transcriptomic layers of information has been yet conducted in FL to our knowledge. Building up a collection of *EZH2* mutant and WT FL samples is of paramount importance to address this unmet clinical need.

The general workflow to identify patients to be included in our retrospective cohort is depicted in ***Fig 4a.*** A total of 160 cases of FL were initially selected for DNA and RNA extraction (elementary patient data are summarized in ***Tab S4a***). Sanger sequencing of *EZH2* exons 16 and 18 allowed identification of 30 (18.8%) mutant EZH2 cases, which is consistent with previously reported rates[10,27–30]. Y646N was the most frequent variant identified in our series (46.7% of mutant cases, ***Tab S4b***). Based on the amount of available material, we then assembled a restricted cohort of 32 patients for further investigation: 21 *EZH2* mutated and 11 *EZH2* WT cases (main patient characteristics are summarized in ***Tab S4c***). All cases displayed the classical immunohistochemical features of FL (co-expression of CD10 and Bcl2) and harbored the t(14;18) translocation, as assessed by FISH and/or karyotyping (one example of a core needle biopsy (P#154) is shown in ***Fig 4b***). Chromatin was extracted from at least one sample for each patient. Sequential data using DRAGON panel (i.e. a in-house tool aiming at the Determination of 571 Relevant Altered Genes in Oncology by NGS), as well as RNA-Seq and H3K27me3 ChIP-Seq were compiled throughout the cohort as shown in ***Fig4c***. Median follow-up was 11.5 years [1-24]. Sex ratio of women to men in our cohort is 2.5:1, while men tend to be slightly more affected (1.2:1) by FL in general [31]. The majority of newly diagnosed FL cases in our institution arise from follow-up of patients initially treated for breast cancer, likely explaining that switch. However, there is no association between the two cancer types nor specific features for the FL cases diagnosed in this manner; therefore, we believe this recruitment bias has no impact on the focus of our work. Given the relatively lower number of mutant cases, samples from patients with mutant FL included in the restricted cohort tended to be, overall, more ancient cases than the WT ones. This explains the higher death rate in the mutant group at the time of the last update of the current study (June 2021). Thus, no inference can be made about the impact on prognosis of the *EZH2* mutation. Moreover, neither stage nor risk stratification at diagnosis are more pejorative in the former group compared with the latter.

NGS sequencing (DRAGON panel) of one DNA sample per patient confirmed all the 21 Sanger-retrieved mutations of *EZH2* (***Fig 4d***). We detected a mean of 7.5 [2-13] pathogenic variants per tumor (i.e. tumor mutational burden). The majority of mutations we identified were missense, 98% of these resulting in a change of protein. The overall mutational landscape of our *EZH2* mutant-enriched cohort was comparable to that of previously published data, with *KMT2D* and *CREBBP* being the most frequently mutated genes and with 29/32 cases displaying ≥ 2 mutations in *KMT2D*, *CREBBP* and/or *EZH2* [29,30]. Of note, DRAGON panel sequencing of two longitudinal samples belonging to patient P#96 and taken 3.3 years apart yielded slightly different results: the earliest sample harbored 3 mutations including one in *BCL2*, while the latest sample harbored a total of 6 mutations without mutation of *BCL2,*

illustrating inter tumoral heterogeneity in FL. Four cases contained multiple mutations in *EZH2*: in addition to mutations at Y646, we detected a splice site mutation inducing no protein changes (1 case), E641D (1 case), and V637A (1 case). Finally, one case carried at the same time 10% Y646F and 22,8% Y646N mutant alleles. The variant validated upon Sanger method was Y646N, with a doubt on the co-existence of Y646F mutant. All *EZH2* mutations were heterozygous, with a median Variant Allelic Frequency (VAF) of 22.9% [7.4 – 57.4]. When adjusted by tumor cellularity (available in 18/21 cases), mean VAF was 46.9% [22.7 – 83.1], suggesting that *EZH2* mutation is a clonal event throughout our cohort. Copy Number Variant (CNV) analysis performed in 29 cases (3 were non-assessable) revealed the presence in at least 3 cases throughout the cohort of 1p cnLOH (copy number neutral loss of heterozygosity, also known as acquired uniparental disomy), 2p+, 6p cnLOH, 7p+, 8q+, 10q-, 12+, 12q+, 16+, 16p cnLOH, 17q+ (***Fig 4e).*** Of note, no amplification of *EZH2* was identified, regardless of the presence of gain in 7p identified in 3 *EZH2* mutant cases. Interestingly, we confirmed the high rate of 1p cnLOH throughout the cohort (14/29), largely overlapping with the cases mutated for tumor suppressor *TNFRSF14* (localized on chromosome 1p) as 10/17 of these cases were also characterized by the aforementioned abnormality. Co-occurrence of 1p cnLOH and mutations in *TNFRSF14* is a well-known mechanism for the loss-of-function of this tumor suppressor and is sometimes reported with worst prognosis in FL [32]. No genomic feature (mutation or CNV) significantly co-occur with *EZH2* mutation in our cohort in regard of WT cases, which is not surprising given the reduced effective of our series.

In summary, we have built a unique longitudinal collection of clinical samples that could be leveraged through further integrative analyses to refine current knowledge on the role of *EZH2* mutation "*in situ* ", in regard of FL genomic background and relapsing clinical pattern.

H3K27me3 redistribution in *EZH2* mutant FL induces repression of both genes involved in immune cross talk and B cell specific enhancers

To investigate the effect of *EZH2* mutation in a tumoral context, we performed integrative H3K27me3 ChIP-Seq/RNA-Seq analyses in a subset of *EZH2* mutant and WT FL samples (**Fig 4a & 4c**). We reasoned that the impact of the mutation on H3K27me3 landscape might not be detectable below a certain threshold for *EZH2*^{Y646} VAF, as other authors were able to define a transcriptomic signature specific to *EZH2* mutant FL only in cases with VAF >17% [27]. Thus, we selected 8 cases with *EZH2*^{Y646} VAF >17% with matching RNA available in addition with 4 WT cases in order to build a discovery set of 12 cases for downstream analyses. Similar to our observations in iMEFs, mutant EZH2 induces a global redistribution of H3K27me3

across the genome with a flattening of the peaks observed in the WT samples alongside with a spreading of the mark in the flanking vicinity of the peaks (***Fig5a, FigS5a***). Of note, this mutant-related pattern does not seem to be more pronounced in cases we the highest VAF compared with those with lower VAF (***FigS5a***). Even though the relative global increase of H3K27me3 in mutant cases is discreet (***Fig5b***), H3K27me3 deposition pattern is strikingly different in these cases and forms a highly enriched "plateau" all along the gene body compared with WT samples. This enrichment contrasts with a depletion of the mark upstream the TSS in the mutant background compared with its WT counterpart (***Fig 5c***).

In the PCA realized upon differential gene expression analysis, PC1 and PC2 both explain less than 15% of the difference between the two groups of patients (***Fig 5d***), which most likely relates to the heterogenous genomic background within the tumors, and potentially the co-mutation of other chromatin modifiers. Consistent with not only our observations in iMEFs but also with other transcriptomic studies in FL [27,28], mutant *EZH2* has a dual effect, with the number of upregulated genes (452 transcripts) falling in the same range of downregulated genes (470 transcripts) (***Fig 5e, Tab S5a***). Our signature nonetheless correctly segregates *EZH2* mutant from WT samples in an independent validation set (***Fig 5f***). However, it did not cluster patients from a published cohort based on *EZH2* mutational status [28]. Interestingly, ontology analyses revealed that genes repressed by mutant EZH2 correspond to a much higher number of retrieved GO terms (123 GO terms, FDR <.05, Panther 15.0) compared with activated genes (37 GO terms), suggesting that repression by the mutant selectively proceeds at PRC2 targets that are relevant for lymphomagenesis whereas gene activation might be less specific. Consistent with previous reports, downregulated genes are mostly involved in immune response related pathways and crosstalk between B cells and non-lymphoid cells (***Fig 5g***). More specifically, 8 sets of genes have been recently identified in tumors from Cγ1-cre mice and shown to be repressed by $Ezh2^{Y641F}$. Out of these sets, 5 are responsible for maintaining B cells in the light zone of the germinal center along with reduced interactions with Follicular Helper T cells. Beyond the species difference, 3/5 of these sets were also enriched in GSEA analysis of downregulated genes from our cohort (FDR<25%, nominal p value <1%, ***Fig 5Sb***) [14]. Equivalently, a list of 258 genes which expression is meant to be critical in the biology of the immune niche of FL were enriched in the target genes of mutant EZH2 in our series (***Fig 5Sc***) [33]. Altogether, these observations made in FL patient samples also converge toward a role for $EZH2^{Y646}$ in reprogramming of the immune niche.

We further addressed the question of the direct impact of H3K27me3 deposition and gene deregulation, focusing separately on both the promoter regions and the gene bodies, given the relatively distinct enrichment pattern of the mark at these two functionally distinct regions upon

EZH2 mutant (***Fig5b***). As for the promoters, the number of regions with gained H3K27me3 is relatively higher in the WT samples compared to the mutant ones but this does not correlate with specific transcriptomic changes (***Fig S5d, upper panel, Fig5h, lower panel***). Conversely, gene bodies decorated with H3K27me3 are, as expected, more numerous in the mutant and the corresponding transcripts tend to be repressed. Thus, the repressive effect of mutant *PRC2* appear *to* relate to the more abundant deposition of H3K27me3 along the gene bodies (***Fig 5h).*** Given the global rewiring of H3K27me3 deposition we observed in all chromatin compartments in *Ezh2*^{Y641} iMEFs, we speculated that its redistribution might impact transcription away from PRC2 canonical regions, and more specifically at enhancer regions. Moreover, the increase of H3K27ac observed in *Ezh2*^{Y641F} iMEFs intrigued us as it could potentially account for the indirect activation of enhancers. B cell-specific enhancers were then defined by the colocalization of H3K27ac and H3K4me1 according to publicly available data. Upon differential binding analysis, we observe a significant enrichment of H3K27me3 at B cells enhancers in *EZH2* mutant samples versus WT, in association with a repression of controlled genes (infered by vicinity with the enhancers) compared with the expression level of a randomly selected equivalent number of genes for which there is no binding of H3K27me3 at the enhancers (***Fig 5i***). Similar differential binding analysis performed on neutrophil specific enhancers shows a higher enrichment in mutant over WT samples reflecting the global spreading of H3K27me3. However, there was no differential expression of the related genes between the two groups, underlying the context specific impact of EZH2 mutant at enhancers in FL (***Fig 5Se***).

Taken together, our results show in human samples the impact of *EZH2*^{Y646} at both chromatin and transcriptomic levels and provide meaningful insights for further investigation about the reprogramming of chromatin beyond H3K27me3 in FL.

Discussion

Beyond H3K27me3: How an unbiased cell line model helps to refine knowledge about the role of PRC2 loss-of-function in cancer

Our integrative analysis of the most frequent alterations of PRC2 encountered in cancer in an isogenic model provides an unbiased comparison of their respective effects on chromatin and transcriptome. We show that the PRC2 loss-of-function effects of the H3.3K27M mutation are milder when compared to those of *Ezh2*-KO, indicating the persistence of residual EZH2 activity in these cells. In parallel to the gradual reduction of H3K27me3 in the different PRC2 loss-of-function contexts, we observe a profound reshuffling of other histone marks. In H3.3K27M and *Ezh2*-KO, H3K27me2 is globally reduced with however a certain retention at Polycomb targets for H3.3K27M. We are planning to perform ChromHMM analysis in *Eed*-KO cells as well: a complete impairment of PRC2 will lead this time to a complete loss of both H3K27me3 and H3K27me2 deposition and probably of H3K27me1. It will be very interesting to determine how this "vacancy" of methylation on the lysine 27 (by far the most abundant post-translational modification of this residue), will affect other histone modifications in particular the ones know to be antagonistic.

The most obvious question is how H3K27 acetylation is affected. While H3K27ac is globally increased in the CUT&RUN experiments, we observed a reduction of this mark at its peaks in both H3.3K27M and *Ezh2*-KO cells as also recently reported in H3K27M mutant neural stem cells by Brien *et al* [34]. Of note, changes in H3K27ac global level confirmation by Western blotting in our cells is ongoing. These changes are particularly interesting as the increase of H3K27ac upon PRC2 loss-of-function has been suggested to constitute an actionable therapeutic target [34,35]. Indeed, acetylated lysine-binding bromodomain proteins, such as BRD2 and BRD4, play a pivotal role in promoting transcription through their recruitment at enhancers. Other authors proposed that chromatin states observed at enhancers in H3K27M glioma cells more likely related to cell state/lineage, and that the mutant did not generate a specific H3K27ac deposition pattern leading to a noticeable sensitivity to bromodomain and extra terminal domain (BET) inhibitors [22]. Alternatively, we suggest that the reported sensitivity of mutant cells to BETi might stem from the pre-existing erosion of H3K27ac: focal depletion of the mark would facilitate its abrogation upon BETi exposure. Moreover, observations made in our model argue against a link between H3.3K27M incorporation and the redistribution of H3K27ac as it has been proposed [34], as various mechanisms of PRC2 impairment lead to similar effect on the active mark.

Retention of H3K27me3 peaks was reported in H3.3K27M glioma mice, prompting the authors to propose using EZH2 inhibitors as a therapeutic strategy in H3K27M midline glioma, whereas previous study showed no impact of Tazemetostat on H3K27M glioma cells growth [18,36]. These observations are contrasting with recent data showing that the sensitivity to EZH2 inhibition in H3.3K27M neural stem cells arises from the specific repression of neurodevelopmental genes [34]. In our hands, no such partial gain-of-function (maintenance-of-function?) of PRC2 upon H3.3K27M is noticed as H3K27me3 shows an unequivocal decrease all over the genome. This discrepancy might reflect the epitranscriptomic features specific to a given cell identity. In order to strengthen our conclusions, the assessment of PRC2 localization to chromatin, using, for instance, Suz12 CUT&RUN-Seq performed in all the genotypes we modelized, would be of particular interest.

The intriguing change-of-function induced by EZH2 mutation: What did we learn? What did we miss?

On the opposite side of the spectrum of cancerous PRC2 alterations, EZH2^{Y641} not only leads to an increased and redistributed deposition of H3K27me3, but also to a major reduction in repressive mark H3K27me2. Overall, H3K27me3 genome-wide distribution in mutant cells follows that of H3K27me2 in WT cells, underpinning the hyperactive nature of the enzyme: Indeed, EZH2^{Y641} seems to result in a communication vessels effect between H3K27me2 and H3K27me3. Nonetheless, partial inhibition of EZH2 in mutant cells revealed that H3K27me3 is not re localized to its usual “residential” loci (sometimes referred to as nucleation sites [25]) as previously reported [15]. Our results thus favor a model where the mutant enzyme is broadly recruited and distributed beyond the narrow nucleation sites, this distribution remaining unchanged upon inhibition. These observations are consistent *w*ith a change-of-function that is irreversible by nature but becomes less potent upon inhibition. They also question what drives the WT enzyme at nucleation sites and, therefore, why this recruitment is modified with the mutant enzyme: if the presence of pre-existing H3K27me does not seem mandatory for PRC2 recruitment, it has been suggested that PRC2 cofactors might participate in the limited recruitment of the complex at its nucleation site through their binding to SUZ12 [37]. However, SUZ12 structure is not altered in B cell lymphoma. One could suggest to investigate the role of PRC2 cofactors in presence of mutant EZH2.

The enzymatic dependency of mutant EZH2 on its WT counterpart has been largely accepted as paradigmatic based on the heterozygous fashion of *EZH2*^{Y646} in B cell lymphoma. Moreover, enzymatic potency studies clearly demonstrated the uneven ability of both mutant and WT alleles to perform H3K27 methylation reaction [23]. However, puzzling *in vivo* data suggested that H3K27me3 global level could still be increased in EZH2$^{Y641/Y641;}$ *Ezh1*-KO background [15].

We demonstrated that not only PRC2-Y641 had a lower, still significant capacity to catalyze H3K27 methylation, but also that PRC2 cofactors played a part in enhancing the reaction. Our results do not question the haploinsufficiency of mutated *Ezh2* observed in FL or DLBCL, they rather underline the complexity in the change-of-function induced by the mutation and open the door, again, to investigations about the potential contribution of PRC2 cofactors to lymphomagenesis *in vivo*.

Understanding the mechanics that underlies the response to EZH2 inhibition has major implications. Indeed, EZH2i Tazemetostat, approved for treatment of FL patients, has a higher potency in *EZH2* mutant background than in the WT tumors. Nonetheless, this latter subset of patients still displays a significant rate of clinical response that question the existence of distinct mechanisms of action of EZH2i in regard of the mutational status of *EZH2*. Previous works that addressed this question in lymphoma cell lines confirmed the cytotoxic effect of EZH2i specific to mutant cells compared with WT cells [23,38,39]. A stronger transcriptomic response (i.e. re expression of more numerous PRC2 targets subsequent to a decrease of H3K27me3 in the presence of EZH2i) in the mutant background was suggested as an explanation for these observations. Response to EZH2i might also strongly relate to cell identity, as WT *EZH2* lymphoma cells seem to be more dependent on BCR activation than mutant *EZH2* cells, illustrating the consequence of EZH2i on B cell maturation [40]. Nonetheless, the high genomic variability among lymphoma cell lines represents a limit to infer a global mechanism explaining the difference of response in regard of the genotype. In contrast, we show that the differential response does not rely on the quantity of responsive genes, but on their nature instead. Indeed, the overlap of responsive genes in both *Ezh2* WT and mutant cells is limited. Moreover, if the response in the WT background seems to correlate with H3K27me3 baseline level around the TSS of responsive genes, similar observation does not hold in Ezh2 mutant cells, where H3K27me3 enrichment is less affected by the EZH2i. We speculate that the response could be mediated by indirect effects on chromatin structure or by the complex regulation of chromatin around enhancers in the mutant cells. In support of this hypothesis, we observed a global increase level of H3K27ac in *Ezh2*^{Y641F} cells, both at peaks observed in the WT cells and in flanking regions. Of note, active chromatin states and enhancer sates appear to gain slightly more H3K27ac than other regions in the mutant background.

Recently, it has been suggested that changes in H3K27ac deposition observed in B cell lymphoma cells specifically impact enhancer-promoter interactions, along with transcriptomic consequences that ensue [41]. Indeed, the authors take advantage of DLBCL WSU-DLCL2 cells to demonstrate that H3K27ac not only display repositioning of the mark at enhancers compared with normal B cells but also participates in oncogenes expression that is reversible

upon histone acetyl transferase inhibitor. Our finding that H3K27ac increase might be a direct consequence of mutant EZH2 leads us to hypothesize that the mutation creates a competition at B cell specific enhancers between those where the effect of increased H3K27me3 prevails over the enhanced deposition of H3K27ac and vice versa. Moreover, as WSU-DLCL2 cells harbor *Ezh2* Y641F mutation, it would be interesting to address this hypothesize in parallel in *Ezh2* WT cells as well. Our isogenic cells might appear to be a model of choice to tackle this question.

We show that redistribution of H3K27me3 in *EZH2* mutant FL does impact B cell specific enhancers, along with expression of related genes that ensues. Altogether, it is tempting to speculate that one of the contributions of the mutation of *EZH2* to lymphomagenesis could be a global reshuffling of enhancers activity consequent to i) an increased deposition of H3K27ac at enhancers ii) the altered balance between H3K27me3 and H3K27me2. Interestingly, an overwhelming majority of FL is characterized by the presence of either *KMT2D* or *CREBBP* mutation or both, which disturbs enhancer activation. The interlaced effects of multiple mutant chromatin modifiers are challenging to interrogate. We believe that our precisely annotated collection of FL samples represents a valid material to leverage for such purpose.

Bridging cell lines and patient samples: the pivotal role of the immune crosstalk

Through our translational integrative study, we establish a link between increased deposition of H3K27me3 over the gene body and gene silencing in *EZH2* mutant FL. However, we did not observe a correlation between gene upregulation in the mutant cells and H3K27me3 pattern as it has been suggested [15]. We hypothesize yet that the global alteration of the chromatin landscape in particular at enhancers might contribute to the mutant phenotype. H3K27ac ChIP-Seq experiments on the same subset of 12 FL we analyzed in our preliminary study (4 *EZH2* WT and 8 *EZH2* mutant cases) is ongoing and we believe the generated data will allow a precise annotation of the enhancers state in both genotypes. The set of genes repressed upon *Ezh2*^{Y641F} mutation, in line with the increase of H3K27me3, was largely enriched in pathways controlling interactions with B cell microenvironment and the immune reaction, which, in regard of previous studies performed in mouse, is consistent with a role for mutant *EZH2* in the modulation of the immune niche in FL [14]. More specifically, the mutation has been proposed to repress key players of the GC exit, thus resulting in the entrapment of FL cells in the light zone where a skew from Tfh to FDC cells operates in the prioritized cross talk between tumor cells and ME [14]. The question of how global effects on the chromatin landscape observed in PRC2 mutant cells can have very specific functional consequences on FL cells remains open. This lineage related-effect triggered by the mutation might stem from numerous factors that determinate the transcriptional state of genes crucial to B cell

differentiation, beyond the sole distribution of H3K27me3. Integrative epitranscriptomic analyses performed in clinical samples, coupled with tissue imaging-based studies, should provide meaningful information about how crosstalk between FL cells and ME is altered *in situ*.

More unexpectedly, we observed the specific upregulation of genes involved in antigenic presentation in *Ezh2*$^{Y641F/WT}$ upon Ezh2 inhibition in iMEF, a model distant from B cells ontogeny. It has been established that *EZH2* mutation triggers MHC deficiency in DLBCL [42]. Thus, restored MHC molecules expression on the cell surface upon EZH2 inhibition provides a rationale for the use of therapies combining EZH2 inhibitors with immunotherapy. The ability of PRC2 to silence antigen presentation pathway is a highly conserved mechanism hijacked by cancer cells that has been shown to promote immune evasion, beyond the scope of B-biology and *EZH2* mutations [26]. Such mechanisms would relate to poised chromatin states at MHC loci observed in multiple cell types. Our results are coherent with that evolutionary conserved specific trait for PRC2. However, in our hands this remain a characteristic feature specific to cells harboring *Ezh2*^{Y641F} mutation. We thus suggest that, in an isogenic model, PRC2 alone is not capable to drive such silencing. A similar mechanism in cells with high reliance on the presence of mutant EZH2 could explain similar findings in DLBCL. The role of the global chromatin environment in this process is then, once more, implicitly questioned.

Our observations in iMEF echo the reported role of mutant CREBBP in suppressing antigen presentation in FL. Indeed, *CREBBP* mutations, that are highly frequent in FL, are long time reported to result in a druggable immune evasion because of aberrantly MHC genes repression [43,44]. In our restricted cohort of FL samples, orthologs included the same GO term were not enriched in downregulated genes upon *EZH2*^{Y646}. Broader investigations about antigenic presentation genes expression level in larger annotated FL cohorts still remain to be conducted. Indeed, it is very likely that both mutated *CREBBP* and *EZH2* converge toward one/some common physiopathological effects in FL cells, including immune evasion: In *Ezh2*^{Y641F} iMEF, MHC genes retain H3K27me3 at their promoter while *Crebbp* deficiency leads to preferential loss of H3K27ac at their enhancers in VavP-Bcl2 mice. The complementarity of these two mechanisms studied within one single model might constitute a meaningful question to address in order to disentangle the influence of chromatin status on oncogenic pathways activation.

Conclusions & perspectives

Epigenetic studies have permitted to establish molecular connections between the genome and cues from its close surroundings. They have unveiled critical mechanisms in developmental biology and have helped answering questions non-entirely solved upon traditional genomic approaches only. In that regard, research about processes that drive and/or trigger cell fate decisions has contributed a great deal to cancer research. Biological questions such as differentiation, control of cell division or maintenance of genome integrity are indeed central to the initiation and progression of cancerous diseases. Discoveries about epigenomic alterations in cancer went a rapid pace and mutations in chromatin modifying genes, identified in a wide spectrum of tumor types, have emerged as enticing therapeutic targets through plethora of preclinical studies. Polycomb genes and proteins (PcG) epitomize this filed of investigations.

PcG have been known for around 80 years. Since then, comprehension of their contribution to physiological processes has been increasing constantly. In parallel, genome-wide analyses have uncovered a variety of mutations affecting the PRC2 machinery in cancers. Given the high tumor type-specificity of these mutations, most studies investigated their functional contribution to cancer development within the context of their respective cell-identity. Deciphering the role of altered PRC2 in a given malignancy might be strongly informative not only about the phenotype of tumor cells but also about phylogenetic and developmental processes specific to their normal counterpart. However, it is much less informative regarding the mechanisms explaining how PRC2 impairs gene expression in cancer or the mechanisms controlling the response to EZH2 inhibitors for instance. This concern formed the setting for my PhD project.

The first aim was then to propose a single isogenic cell line model recapitulating the most common alterations of PRC2 encountered in cancer, in order to dissect their relative impact on chromatin environment and transcriptome.

Our analyses revealed that, in case of partial loss-of-function of PRC2, the consequences on the global level of H3K27me3 and the repercussion of such loss on several other histone marks are uneven, depending on the type of alteration of PRC2. Moreover, we showed, in our purely mechanistic model, the unambiguous loss-of-function of PRC2 that is imputable to H3.3K27M mutation, in contrast to what has been proposed in other models. We demonstrated that Ezh2^{Y641} mutation deeply reshuffled not only H3K27me3, but also H3K27me2 deposition landscape, relating to both gain and change-of-function of PRC2 enzymatic activity. We also unveiled a role for the cofactors of PRC2 to enhancing the ability for PRC2-EZH2^{Y641F} to

catalyze the monomethylation of H3K27 in vitro. Finally, we proposed that the differential response to EZH2 inhibitor observed between Ezh2 ^{Y641} and Ezh2 WT cells stems from a different nature of the responsive genes according to the genotype.

Our results may have several impacts and broaden perspectives for upcoming investigations. First, as they allowed us to identify the substantial changes of chromatin composition upon PRC2 alterations, they underline the relevance of investigating how each reprogrammed histone mark influence gene expression – may it be the consequence of a stochastic deposition or in correlation with specific chromatin state. Thus, we will extend our explorations about the precise impact of H3K27ac and H3K27me2 redistribution in regard of PRC2 both gain and loss-of-function. We will undertake the effort to characterize more precisely the chromatin environment of the genes which expression is affected by these mutations. Alternatively, our isogenic model represents an opportunity to study the response to a wide range of chromatin targeting compounds, beyond the scope of EZH2 inhibitors.

Our results stress the relevant complementarity of both a mechanistic approach on the one hand, and a translational approach in the other hand. Still, translating bench science to the clinical setting remains a significant hurdle. One remarkable example of such success, however, is the story of EZH2 mutation in follicular lymphoma: A decade separates the first (and incidental) report of the mutation in a patient with FL and the FDA approval for Tazemetostat, an EZH2 inhibitor highly potent in patient with EZH2 mutant FL. Beyond this striking clinical accomplishment, understanding the precise role played by the mutation within the complex biology of FL cells and their environment remains an ongoing effort. How does mutant EZH2 cooperate with other chromatin modifiers to maintain GC B cells inside of the GC and to promote the reprogramming of the microenvironment, or to drive response/resistance to EZH2 inhibitor? What is the impact of the mutation on H3K27me2 and H3K27ac redistribution and the gene expression that ensues? How does EZH2 mutation exert its role throughout the relapsing course of the disease? We reasoned than capitalizing on tumor samples would be particularly relevant to tackle these questions.

As a pathologist, I have had the opportunity to notice that examination of tumor sections is a highly valuable source of information about tumor biology: morphological features of tumor cells, composition of the microenvironment, intratumor heterogeneity are examples of hints for drawing hypotheses about how molecular mechanisms can drive oncogenesis. The rising development of technologies allowing the investigation of epitranscriptome directly form patient samples represents a unique opportunity in order to challenge these hypotheses and confront results obtained in fundamental models. In that regard, we believe that some of our findings

made in the iMEF model about the biology of Ezh2 ^{Y641} could be exploited in our cohort of FL patient samples.

This is why the second aim of this work was to create a translational tool relevant for in depth characterization of EZH2 mutant FL in regard to its WT counterpart.

The clinico-biological collection we constituted is, to our knowledge, unique in its kind. Even though still largely underexploited within the span of this PhD project because of time constrains, we were able to provide elementary information about the mutational environment of our samples, which will be determinant as we are projecting to interrogate the contribution of combined chromatin modifying genes mutations along with EZH2 mutation. In order to more accurately address this question, we are also planning, in parallel, to introduce mutations of Crebbp and/or Kmt2d in Ezh2 mutant iMEF, and subsequently conduct CUT&RUN-Seq/RNA-Seq analyses. We objectified the consequences of the mutation on H3K27me3 distribution and its impact on gene expression: Our results suggest that mutant EZH2 might trigger enhancer activation. Based on the availability of the chromatin we were able to retrieve from either fresh frozen or FFPE samples, we are then planning to further realize H3K27ac and H3K27me2 ChIP-Seq. Moreover, combined histone PTM and mass spectrometry analyses carried out on FFPE samples have been proposed recently [41]. This approach might find an appropriate application in the context of FL. Longitudinal samples will give us an opportunity to address the question of the evolutionary behavior of H3K27me3 in regard of the clonality pattern of EZH2 mutation. Finally, our combined ChIP-Seq/RNA-Seq approach of FL tumor samples rose the interest of two pharmaceutical companies working on the development of EZH2 inhibitors. Being able to monitor the effects of the treatment in tumor tissue would be paramount toward a better comprehension of the mechanisms of action /resistance of the drug.

In total, this PhD project allowed, through the mean of an original dual approach, to broaden knowledge about how PRC2 alterations encountered in cancer can reshape chromatin landscape and, thus, modulate gene expression. Given the high occurrence of recurrent mutations in chromatin modifying genes, in FL in particular, and in cancer in general, studying epitranscriptome from clinically annotated tissue material will provide determinant information in regard of tumor progression and/or response to treatment.

Material & Methods

Cells

Cell lines were derived from Ezh2 flox/flox; 4-hydroxytamoxifen-inducible ROSA26 Cre (CreERT2) mouse embryonic fibroblasts (mEFs) immortalized by overexpression of c-Myc. Cells were grown in Dulbecco's Modified Eagle Medium (DMEM)- F12 + glutamine (Gibco), 1% N2 (Gibco), 1% P/S supplemented with 10% FBS, 1mM L glutamine, 100 mM nonessential amino acids and 1% Penn/Strep at 37 degrees Celsius with 5% CO2 and 98% humidity.

Introduction of a heterozygous Ezh2$^{Y641F/WT}$ mutation

CRISPR-Cas9-mediated cut of target DNA was used along with a repair template containing the Ezh2-Y641F mutation as previously described [1]. In addition to introducing the point mutation, this procedure eliminates the 5' LoxP site, rendering the Ezh2-mutant allele insensitive to Cre-mediated excision. In brief, the right arm of the locus targeted by Cas9 partly overlaps with Ezh2 exon 16 in which an adenine is replaced by a thymine, resulting in the mutation of the tyrosine Y641 in a phenylalanine (Y641F). In between the two homology arms, a selection cassette flanked by FRT sequences is introduced, thus allowing its excision upon transfection with flippase (FRT-FLP recombination). Upon excision, a FRT sequence remains in the intron. Within the cassette, a splicing acceptor placed prior to the T2A sequence allows the correct expression of the cassette until the polyA signal. The T2A sequence enables the translation of two proteins from a unique coding sequence. A hygromycin resistance gene, was placed in between T2A and polyA signal, allowing resistant cells to be selected for further analysis. Sanger sequencing validated the substitution of adenine to thymine in codon 641 resulting in the Y641F mutation.

Conditional deletion of Ezh2 wild-type allele

Deletion of the SET catalytic domain of *Ezh2* on the WT allele was induced by supplementing culture medium with 1 μM 4-hydroxytamoxifen (TAM) for 7 days. Addition of an equivalent volume of ethanol (EtOH) to the medium was used as a control. Cells were extracted after the 7 days in tandem. Sanger sequencing validated the deletion of *Ezh2*.

Generation of KO and H3.3K27M mutant cell lines

Generation of KO/mutant mEFs was performed using CRISPR/Cas9 technology as well. In brief, a stop cassette containing an antibiotic resistance gene followed by polyadenylation sequence was inserted into early exons of target genes by homologous recombination. After antibiotic selection, clones were genotyped and complete KO was validated by western blot.

Similar technology was used in mEFs containing *Ezh2*$^{Y641F/-}$ mutation in order to introduce *Ezh1* point mutation.

Chemical inhibition of PRC2 activity

Cells were treated with 0.125 µM, 0.25 µM, or 2µM of UNC1999 or mock control (UNC2400) for 10 days. Culture medium was renewed every 2 or 3 days. Cells were passaged constantly when reached 80% of confluence. Cells were harvest for RNA and protein extraction or CUT&RUN after 10 days of treatment.

RNA extraction, RNA -Seq analysis

All cells were grown in tandem and RNA was extracted 3 days after passaging in 200,000 cells in 25cm^2 flasks per condition per clone. Once 70-80% confluency was obtained, 10^6 cells were passed a 2nd time into 75cm^2 flasks. Nuclei were extracted 5 days later at 70-80 % confluency (cf below).

Total was isolated using Trizol-Chloroform extraction and isopropanol precipitation. After RNA was extracted, a gel electrophoresis analysis was performed in order to verify the quality of the RNA. cDNA was further obtained using the High Capacity cDNA RT kit (Applied Biosystems).

50 bp single ends reads were generated using HiSeq2500 (Illumina) with the Rapid Run mode. Raw reads were trimmed for adapters with cutadapt (1.12) using the Trim Galore! (0.4.4) wrapper (default settings) and subsequently mapped to the complete mouse rRNA sequence with Bowtie2 (2.2.9). Reads that did not map to rRNA were then mapped with STAR (2.5.2b) to the full reference genome (mm10) using the following parameters: -outSAMtyupe BAM SortedByCoordinate -runMode alignReads -outFilterMismatchNmax 6 –outFilterMultimapNmax 20 -outSAMmultNmax 20 -outSAMpriomaryFlag OneBrestScore. Gene counts were generated using STAR -quant_mode (uniquely mapped, properly paired reads that overlap the exon boundaries of each gene). The resulting count table was normalized using DESeq2 method and log2-transformed for all downstream analyses.

Gene Ontology (GO) analysis

We first identified genes differentially expressed between WT and *Ezh2*$^{Y641F/WT}$ iMEFs (both up- and down-regulated). We also defined genes significantly up-regulated in either WT or *Ezh2*$^{Y641F/WT}$ iMEFs upon treatment with 2 mM UNC1999. The 4 lists of deregulated genes were then submitted to http://geneontology.org/ for identification of overrepresented GO terms [2–4]. To build the bubble plots in Figures 3b,3g, S3b-d we retrieved the enrichment and significance values for all GO terms and discarded those with a fold enrichment >100 or <0.01, corresponding to GO terms encompassing a very small number of genes or pathways with no matches in our lists of differentially expressed genes. We then manually annotated the significantly overrepresented GO terms (fold enrichment >= 2, FDR < 0.05) into broad categories (e.g. cell communication, cell migration and adhesion, or metabolism). Finally, we computed the median z-score for the significantly deregulated genes overlapping each GO term. Bubble plots were generated with the ggplot2 R library [5].

Preparation of nuclear extract and immunoblotting

Cells were incubated with buffer A (10mM Hepes at pH 7.9, 2.5 mM MgCL2, 0.25 M sucrose, 0.1% NP40, 0.5 mM DTT, 0.5 mM Aprotinin, 0.5 mM Leupeptin, 0.5 mM Pepstatin, 1 mM PMSF) for 10 min at 4°C. After centrifugation, the pellet was then resuspended in buffer B (25 mM Hepes at pH 7.9, 1.5 mM MgCl2, 700 mM NaCl, 0.5 mM DTT, 0.5 mM Aprotonin, 0.5 mM Leupeptin, 0.5 mM Pepstatin, 1mMPMSF, 0.1 mM EDTA, and 20% glycerol) and incubated 10 min on ice. Nuclei were then sonicated for 45 sec with 10% amplitude, then centrifuged at 14,000rpm for 15 min at 4°C. After Bradford quantification, nuclear extracts were acetone-precipitated. All samples were mixed with loading buffer containing SDS and β-mercaptoethanol and run on homemade 15% or commercial 4-15% gradient acrylamide gel (Bio-Rad). Semi-dry transfer was performed on a Trans-Blot Turbo transfer system (Bio-Rad). Correct transfer was verified by Ponceau staining. The following primary antibodies were used: anti-EZH2 (homemade; 1/3000), anti-H3K27me3 (Cell-Signaling; 1/3000), anti-H3K27me2 (Active Motif; 1/5000), anti-H3K27me1 (Active-Motif; 1/3000), anti-EED (homemade; 1/2000), anti-HK27M (Millipore; 1/3000), Anti-H4 (Active-Motif; 1/3000). Starbright Blue 700 (BioRad; 1/5000) fluorescent secondary antibody and HRP (HorseRadish-Peroxidasee; 1/5000) secondary antibodies were used subsequently. Anti-H4 was used to verify that equal protein levels were used for each sample. Homemade antibodies are described elsewhere [6], all the antibodies used in this study are listed in the table at the end of this section.

Western Blot analysis of protein extracts was performed by StarBright Blue 700 fluorescent secondary antibodies (Biorad) and DyLight 800 secondary antibody (Biorad). Imaging was carried out by ChemiDoc System (Biorad). Protein levels and background noise were measured using ImageJ software. Background noise was subsequently subtracted and protein measurements were calculated in comparison to wild-type expression levels (sample/wild-type protein expression level). Results were then normalized to the amount of protein added in each well using the measurements of anti-histone 4 (H4) on the same membrane.

CUT&RUN

CUT&RUN was performed as previously described with minor modifications [7]. In brief, 1 million cells were pelleted at 600 g for 3 min at RT. After washing twice with 1 mL of wash buffer (20 mM HEPES pH 7.5, 150 mM NaCl, 0.5 mM spermidine (Sigma) and protease inhibitors), cells were resuspended in wash buffer and ready for binding with beads. 10 µl of Convanavalin A beads (Bang Laboratories) was washed once with 1 mL binding buffer (20 mM HEPES pH 7.9, 10 mM KCl, 1 mM $CaCl_2$ and 1 mM $MnCl_2$) and placed on magnet stand to remove the liquid. 10 µl of binding buffer was used to resuspend the beads then the slurry was transferred to cells and incubated for 10 min at RT with rotation. After brief spin-down, tubes were placed on magnet to quickly withdraw the liquid. 50 µl of antibody buffer (wash buffer supplemented with 0.1 % digitonin (Millipore), 2 mM EDTA and 1:100 dilution of antibody of interest) was pipetted and cells were incubated for 10 min at RT with mild agitation. Permeabilized cells were decanted carefully and washed once with 1 mL dig-wash buffer (0.1 % digitonin in wash buffer). A secondary rabbit anti-mouse antibody (ab6709, abcam) binding step was carried out if the host species of primary antibodies are mouse. 50 µl of pA-MNase in dig-wash buffer (final concentration of 700 ng/ mL) was incubated with cells for 10 min at RT with agitation. After 2 washes with 1 mL dig-wash buffer, beads were resuspended with 100 µl dig-wash buffer and placed on heat block immersed in wet ice to chill down to 0 °C. 2 µl of 100 mM $CaCl_2$ was

added to activate pA-MNase and incubated on heat block for 30 min. 100 µl of 2 × stop buffer (340 mM NaCl, 20 mM EDTA, 4 mM EGTA, 0.02 % digitonin, 1:200 RNase A, glycogen (50 mg/ mL) and heterologous spike-in DNA (extracted from *Drosophila* S2 cells) (2 pg/ml) was added to quench pA-MNAse, and fragments were released by 10 min incubation at 37 °C with rotation. After centrifugation at 14000 g for 5 min at 4 °C, DNA fragments were recovered by NucleoSpin (Macherey Nagel) or phenol-chloroform purification. Library was prepared by Accel-NGS 2S plus DNA library Kits (Swift Biosciences) for Illumina barcoded system with 16 cycles for amplification. After post-library size selection, library size distribution and concentration were validated by TapeStation 4200 (Agilent). Libraries were sequenced as paired-ended 100bp reads on Illumina Novaseq sequencer.

CUT&RUN-Seq data analysis

Reads were mapped to the mouse reference genome (GRCh37/hg19) with Bowtie2 using default parameters. Aligned reads were sorted by SAM tools. PCR duplicates were removed with Picard Tools MarkDuplicates (https://github.com/bioinfo-pf-curie/ChIP-seq). Generated BAM files were filtered to exclude common artifact regions. (artefact regions: https://github.com/Boyle-Lab/Blacklist/tree/master/lists). Biological replicates were merged with MergeSamFiles for downstream analysis. Reads were counted in bins of length 50, normalized on both library size (RPKM) and scaling factor calculated based on the ratio of heterologous spike-in DNA (dm6 genome) out of mm10 aligned reads, along with -extendedReads normalization, and converted to bigWig format using bamCoverage (v3.3.2.0.0).
H3K27me3 peaks were called with MACS2 (v 2.1.1.20160309.6) with default parameters. For control files, Igg.bam files are used. Minimum FDR (q-value) cutoff for peak detection is modified in each cell line probed. The FDR cutoff was set as 0.01 for H3.3K27M cells and 0.0001 for *Ezh2*$^{Y641F/-}$ cells. Windows of 4 kb centered on transcription start sites (TSSs) were defined as promoter regions. Metaplot and heatmap analyses were performed using deepTools (v3.3.2.0.0): RPKM normalized log2 ratio between treated files from compared conditions were calculated by bamCompare. Matrix was prepared by computeMatrix (v3.3.2.0.0) for metaplot and heatmap visualization. Count table for differential binding analysis was gerenated by FeatureCounts (Subrage package 2.0.2).
ChromHMM (v1.22) and ChromDiff analyses were perform using concatenation of WT and H3.3K27M, *Eed*KO, *Ezh2*KO and *Ezh2*$^{Y641F/WT}$ sequencing files. Biological replicates of CUT&RUN-seq for H3K4me3, H3K27ac, H2Aub, H3K36me3, H3K27me2 and H3K27me3 were fed to the algorithm for state emission following ChromDiff differential analysis [8]. Only consistent transitions in both replicates were considered.

HMT assays

HMT assays were performed as described previously [6]. Briefly, the reaction containing 500ng of PRC2-EZH2 WT or EZH2 Y641F, 1 µg of substrates (recombinant nucleosomes), and 0.2 M DTT was incubated in methylation reaction buffer (50mM Tris-HCl pH 8.5, 2.5mM MgCl2) in presence of 3H--SAM at 30°C for 15 min or 30 min Reactions were stopped by boiling 5 min in SDS Laemmli buffer, run on acrylamide gels and transferred on PVDF membranes. When added to the reaction, peptides JARID2, AEBP2 are at 10--50µM concentrations.

Nucleosomes were generated by salt dialysis. H3K27me3 nucleosomes were generated as described in Voigt *et al* [9].

Patients

Clinical data

All tissue samples were fully anonymized before processing and sequencing. Study approval was first provided by institutional review board and local ethical committee (Groupe Thématique de Travail – Hematology section, Institut Curie), under project ID BS#2014-311 when written informed patient consent was available, or else by the Comité de Protection des Personnes Sud Méditerrannée I, under project ID#RCB2019-A000620-57 when written consent was not available due to death or lost to follow-up. RNA-Seq raw data from patients included in the PRIMA trial (NCT00140582) were courtesy of Drs. Huet, Tesson and Salles (Hospices Civils de Lyon, Pierre-Bénite; INSERM1052, Université de Lyon; Carnot Calym, Pierre-Bénite; France).

Biological samples

All patients with follicular lymphoma were recruited through Institut Curie (Paris, France) based on frozen sample availability and clinical history at the Biological Resource Center. In total, 160 patients were diagnosed and/or treated at Institut Curie for grade 1/2/3a FL between 1988 and 2017. Every patient had at least 1 Formalin Fixed Parafin Embedded (FFPE) and fresh frozen sample stored, 106 had multiple time points samples with either FFPE only or FFPE and frozen samples available. For 160 patients, DNA was extracted from frozen samples after diagnostic confirmation and tumor cellularity during data collection phase. RNA was extracted from 126 frozen samples.

DNA extraction & EZH2 exons 16 and 18 sequencing

20 mg of fresh-frozen tissue from 160 cases were digested with 600 µL of lysis buffer (TrisHCl 50nM, EDTA 50 mM, NaCl 10 mM, SDS 1%), then with 12 µL of proteinase K (PK) before incubation at 55°C overnight. 6 µL of RNase were added before incubation at 37°C for 1 hour. DNA was isolated by phenol-chloroform extraction using standard procedures. The concentration of DNA was assessed via optic density (spectrophotometer Nanodrop, Thermofischer) and 260/280 and 260/230 ratios were calculated. The quality of the DNA was assessed by 0.8% agarose gel electrophoresis.

PCR primers were designed to amplify the 16 and 18 exons and flanking intronic sequences of *EZH2*. Primers were designed using the Primer3 software and NCBI tool Primer-BLAST (https://www.ncbi.nlm.nih.gov/tools/primer-blast/). The program was configured to design primers for PCR products approximately 500bp in length (exon 16: sequence of the forward primer is 5'- TAATGTTCATAGCCATTCTCAGCAG-3', sequence of the reverse primer is 3'-CACAATCCAGTTACTAAGCATGCAA-5', exon 18: sequence of the forward primer is 5'-GCTCTCTTGGCAAAAATACCTATCC-3', sequence of the reverse primer is 3'-GCTTTTGAGTCAGATAACCATCTTG-5'). PCR of genomic DNA was carried out using 100 ng of genomic DNA with (20ng/µl) was added 2.5µl primers (10 µM), 0.5µl dNTPs (10mM each), 5 µl Buffer GoTaq (Promega), 0.2µl GoTaq Enzyme (1.25U/reaction, Promega), 2 µl MgCl2

(25 nM) and 9.8µl sterile water. Cycling was performed on a Bio-Rad T100 thermal cycler. Following an initial denaturation step of heating to 95°C for 2 minutes, were 35 cycles of denaturation at 95°C for 60 seconds, annealing at 53°C for 30 seconds for exon 16 or at 51°C for 30 seconds for exon 18 and extension at 72°C for 60 seconds and a final extension step at 72oC for 5 minutes. PCR products were evaluated by electrophoresis then further processed with the PCR clean-up Kit (Marchery-Nagel) according to the the manufacturer's instructions. PCR product sequencing was carried out by Eurofins Genomics, Germany.

RNA extraction & RNA-Seq, analysis

126 samples were extracted using miRNeasy mini kit (Qiagen, CA, USA) according to the manufacturer's instructions: In brief, 20 mg of fresh-frozen thawed tissue samples were transferred to 500 µL QIAzol-filled stainless steel bead tubes. Samples were homogenized in mixer mill Retsch MM400 for 2 minutes at 30 Hz and then transferred and briefly centrifuged in a phase lock gel heavy (PLGH) tube before adding another 200 µL of QIAzol. After 5 mn incubation at RT, 140 µl of chloroform were added and the samples were centrifuged for 10 minutes at 14000 g, at 4°C. The RNA was precipitated with 525 µl 100% ethanol at RT and then briefly centrifuged. 500 µL of the samples were transferred to a RNeasy Mini Spin Column, centrifuged and washed once with adequate buffers. After elution, RNA concentration was assessed *via* Nanodrop, and 260/280, 260/230 and 28S/18S ratios were calculated. The quality of the RNA was assessed on automated Tapestation 4200 platform (Agilent) according to the manufacturer's instructions. RNA Integrity Numbers (RINs) were used to evaluate the integrity of the RNA samples with > 7.0 considered intact and < 7.0 considered degraded.

100 bp pair ends reads were generated using HiSeq2500 (Illumina) with the TruSeq Stranded mRNA mode. Reads trimming, pairing and mapping were carried out as described previously on the reference human genome (GRch37/hg19), as well as differential gene expression analyses. Gene ontology studies were performed using online Tools ShinyGO (v0.60), Reactome Pathway Browser (v3.7) and GSEA (v4.1.0).

ChIP & ChIP-Seq analysis

For fresh frozen follicular lymphoma samples, approximatively 30 mg of tissue were pulverized in liquid nitrogen, cross-linked with formaldehyde 1% for 8 minutes and treated with 25 µM glycin in order to stop cross-linking reaction for 5 minutes. After 2 washes in PBS, the pellet was incubated in two lysis buffers (LB1: 50mM Hepes-KOH pH7.5, 140 mM NaCl, 1mM EDTA pH8, 10% glycerol, 0.5% NP-40, 0.25% Triton X100. LB2: 200 mM NaCl, 1mM EDTA pH8, 0.5mM EGTA pH8, 10mM Tris-HCl pH8) and then suspended in a sonication buffer (1mM EDTA pH8, 0.5mM EGTA pH8, 10mM Tris-HCl pH8, 100 mM NaCl, 0.1% Na-Deoxycholate, 0.5% N-lauroyl Sarcosine). Each sample was processed using Covaris S220 sonicator at peak power 175, duty factor 10 cycle/burst 200, during 4 minutes.

For FFPE lymphoma samples, chromatin extraction was performed as previously described by Cejas *at al*. In brief, 8 sections of 10 µm thick per sample were deparaffined in xylene 3 times then progressively rehydrated in ethanol decreasing concentration solutions. Samples were then incubated at 40°C for 1 hour before treatment of 5 minutes with PK. Reaction was stopped with Protease inhibitor cocktail containing 1:100 PMSF, 1:000 aprotinin, leupeptin,

and pepstatin A. Each sample was then sonicated using Covaris S200 sonicator at peak power 240, duty factor 20, cycle/burst 200, during 40 minutes.

The sonicated samples (5 µg chromatin/sample approximatively) were cleared during 2 hours with non-coupled magnetic beads (Dynabeads Protein G, Invitrogen), then further mixed with 5 µL antibody (anti-H3K27me3, CST) and incubated overnight at 4°C on rotator. Complexes were pulled down using Protein A Dynabeads (ThermoFisher Scientific), washed with RIPA buffer and decross-linked, and genomic DNA was isolated using standard phenol-chloroform procedure.

Libraries preparation, validation and sequencing followed the same procedure as described above for CUT&RUN analyses, at the exception of the genome of reference (GRch37/hg19) and the absence of spike-in normalization. Data analyses was performed with the same tools as described above.

DRAGON panel

The design of the NGS panel called DRAGON (Determination of 571 Relevant Altered Genes in Oncology by NGS) has been developed specifically for the molecular analysis of tumors in Institut Curie, unit of Genetics (Paris, France). It is composed of 571 genes of interest in oncology from a diagnosis, prognosis and molecular therapy points of view. The nucleotide sequence (variants) as well as the number of copies (deletion and focal amplification) are explored for all of these 571 genes. The panel also includes 86 microsatellites sequences to access the MSI (microsatellite instability) status. Finally, the coding sequence size of the panel exceeds 1.5 Mb and thus allow to assess the Tumor Mutational Burden (TMB) as reliably as with whole exome sequencing (Chalmers et al., Genome Medecine, 2017).
10 ng of input DNA are processed with SureSelect XT-HS library preparation kit that incorporates molecular barcodes (UMIs) to detect variations with very low allelic ratios and effectively eliminate background noise. Total panel size is 2.7 MB. Sequencing is carried out on Illumina Novaseq sequencer with an average depth of 2000X and a minimum depth of 300X.
Bioinformatics pipeline for DRAGON panel is as follows: quality controls, identity controls (based on the polymorphism clustering present in the design as well as on 37 tri-allelic nucleotide polymorphisms to detect inter-sample contaminations, variant calling using Varscan2 (v2.4.3), calculation of TMB based on the number of non-synonymous variant coding per Mb. Using the backbone included in the design, the copy number profile for each tumor is estimated using Facets (v0.5.1) with the median coverage as control. The ploidy, estimated cellularity, and the LOH can also be evaluated by this method.

Somatic mutation calling and annotation

All of the variants were filtered through the COSMIC Cancer Gene Census (CGC) v86 (cancer.sanger.uk). Mutations affecting these putative driver genes were annotated as driver mutations if they passed the following filters:

- The mutation type must have been validated in the CGC for the affected gene
- The mutation must be detected with an allelic frequency >5% in the sample, with a coverage >x100, GnomeAD all must be <0.5%

- We selected frameshift ins and del, as well as stopgain mutation type. Intronic mutation were removed.

Genes identified as recurrently mutated in two large studies of FL were systematically screened [10,11].

Data availability

RNAseq, CUT&RUNseq and ChIPseq data for this study are available upon request and will be deposited in the Gene Expression Omnibus at time of publication.

List of antibodies used in the study

Antibody	Host	Application	Source	Clone/identifier
Ezh2	rabbit polyclonal	WB	home made	N/A
H3K27me3	rabbit monoclonal	WB, C&R,ChIP	CST	C36B11
H3K27me2	mouse monoclonal	WB, C&R	Active Motif	324
H3K27me1	mouse monoclonal	WB	Active Motif	321
EED	rabbit polyclonal	WB	home made	N/A
H3.3K27M	rabbit monoclonal	WB	Millipore	RM192
H2Aub	rabbit monoclonal	C&R	CST	8240S
H3K27ac	rabbit polyclonal	C&R	Abcam	Ab4729
H3K4me3	rabbit monoclonal	C&R	CST	C42D8
H3K36me3	rabbit polyclonal	C&R	Abcam	Ab9050
H4	rabbit polyclonal	WB	Active Motif	AB_2636967

WB Western blot, C&R CUT&RUN, ChIP Chromatin Immunoprecipitation, CST Cell Signaling Technologies

1. Wassef M, Luscan A, Battistella A, Le Corre S, Li H, Wallace MR, et al. Versatile and precise gene-targeting strategies for functional studies in mammalian cell lines. Methods. 2017;121-122:45-54.

2. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, et al. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. Nat Genet. 2000;25(1):25-9.

3. Gene Ontology Consortium. The Gene Ontology resource: enriching a GOld mine. Nucleic Acids Res. 2021;49(D1):D325-34.

4. Mi H, Muruganujan A, Ebert D, Huang X, Thomas PD. PANTHER version 14: more genomes, a new PANTHER GO-slim and improvements in enrichment analysis tools. Nucleic Acids Res. 2019;47(D1):D419-26.

5. Wickham H. ggplot2: Elegant Graphics for Data Analysis [Internet]. New York: Springer-Verlag; 2009. (Use R!). Disponible sur: https://www.springer.com/gp/book/9780387981413

6. Margueron R, Justin N, Ohno K, Sharpe ML, Son J, Drury WJ, et al. Role of the polycomb protein Eed in the propagation of repressive histone marks. Nature. 2009;461(7265):762-7.

7. Skene PJ, Henikoff JG, Henikoff S. Targeted in situ genome-wide profiling with high efficiency for low cell numbers. Nat Protoc. 2018;13(5):1006-19.

8. Ernst J, Kellis M. Chromatin state discovery and genome annotation with ChromHMM. Nat Protoc. 2017;12(12):2478-92.

9. Voigt P, LeRoy G, Drury WJ, Zee BM, Son J, Beck DB, et al. Asymmetrically modified nucleosomes. Cell. 2012;151(1):181-93.

10. Green MR, Gentles AJ, Nair RV, Irish JM, Kihira S, Liu CL, et al. Hierarchy in somatic mutations arising during genomic evolution and progression of follicular lymphoma. Blood. 2013;121(9):1604-11.

11. Okosun J, Bödör C, Wang J, Araf S, Yang C-Y, Pan C, et al. Integrated genomic analysis identifies recurrent mutations and evolution patterns driving the initiation and progression of follicular lymphoma. Nat Genet. 2014;46(2):176-81.

References

1. Vogelstein, B. *et al.* Cancer genome landscapes. *Science* 339, 1546–1558 (2013).

2. Zhao, S., Allis, C. D. & Wang, G. G. The language of chromatin modification in human cancers. *Nature Reviews Cancer* 21, 413–430 (2021).

3. Bracken, A. P. *et al.* EZH2 is downstream of the pRB-E2F pathway, essential for proliferation and amplified in cancer. *EMBO J* 22, 5323–5335 (2003).

4. Wassef, M. & Margueron, R. The Multiple Facets of PRC2 Alterations in Cancers. *J Mol Biol* 429, 1978–1993 (2017).

5. De Raedt, T. *et al.* PRC2 loss amplifies Ras-driven transcription and confers sensitivity to BRD4-based therapies. *Nature* 514, 247–251 (2014).

6. Kaito, S. & Iwama, A. Pathogenic Impacts of Dysregulated Polycomb Repressive Complex Function in Hematological Malignancies. *Int J Mol Sci* 22, (2020).

7. Krug, B., Harutyunyan, A. S., Deshmukh, S. & Jabado, N. Polycomb repressive complex 2 in the driver's seat of childhood and young adult brain tumours. *Trends Cell Biol* (2021) doi:10.1016/j.tcb.2021.05.006.

8. Jain, S. U. *et al.* PFA ependymoma-associated protein EZHIP inhibits PRC2 activity through a H3 K27M-like mechanism. *Nat Commun* 10, 2146 (2019).

9. 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).

10. Morin, R. D. *et al.* Frequent mutation of histone-modifying genes in non-Hodgkin lymphoma. *Nature* 476, 298–303 (2011).

11. Lunning, M. A. & Green, M. R. Mutation of chromatin modifiers; an emerging hallmark of germinal center B-cell lymphomas. *Blood Cancer J* 5, e361 (2015).

12. Caganova, M. *et al.* Germinal center dysregulation by histone methyltransferase EZH2 promotes lymphomagenesis. *J Clin Invest* 123, 5009–5022 (2013).

13. Béguelin, W. *et al.* EZH2 is required for germinal center formation and somatic EZH2 mutations promote lymphoid transformation. *Cancer Cell* 23, 677–692 (2013).

14. Béguelin, W. *et al.* Mutant EZH2 Induces a Pre-malignant Lymphoma Niche by Reprogramming the Immune Response. *Cancer Cell* 37, 655-673.e11 (2020).

15. 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).

16. Morschhauser, F. *et al.* Tazemetostat for patients with relapsed or refractory follicular lymphoma: an open-label, single-arm, multicentre, phase 2 trial. *Lancet Oncol* 21, 1433–1442 (2020).

17. Wassef, M., Michaud, A. & Margueron, R. Association between EZH2 expression, silencing of tumor suppressors and disease outcome in solid tumors. *Cell Cycle* 15, 2256–2262 (2016).

18. Mohammad, F. *et al.* EZH2 is a potential therapeutic target for H3K27M-mutant pediatric gliomas. *Nat Med* 23, 483–492 (2017).

19. Ernst, J. & Kellis, M. Chromatin state discovery and genome annotation with ChromHMM. *Nat Protoc* 12, 2478–2492 (2017).

20. Harutyunyan, A. S. *et al.* H3K27M induces defective chromatin spread of PRC2-mediated repressive H3K27me2/me3 and is essential for glioma tumorigenesis. *Nat Commun* 10, 1262 (2019).

21. Ragazzini, R. *et al.* EZHIP constrains Polycomb Repressive Complex 2 activity in germ cells. *Nat Commun* 10, 3858 (2019).

22. Krug, B. *et al.* Pervasive H3K27 Acetylation Leads to ERV Expression and a Therapeutic Vulnerability in H3K27M Gliomas. *Cancer Cell* 35, 782-797.e8 (2019).

23. 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 U S A* 107, 20980–20985 (2010).

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

25. Oksuz, O. *et al.* Capturing the Onset of PRC2-Mediated Repressive Domain Formation. *Mol Cell* 70, 1149-1162.e5 (2018).

26. Burr, M. L. *et al.* An Evolutionarily Conserved Function of Polycomb Silences the MHC Class I Antigen Presentation Pathway and Enables Immune Evasion in Cancer. *Cancer Cell* 36, 385-401.e8 (2019).

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

28. Huet, S. *et al.* EZH2 alterations in follicular lymphoma: biological and clinical correlations. *Blood Cancer J* 7, e555 (2017).

29. Green, M. R. *et al.* Mutations in early follicular lymphoma progenitors are associated with suppressed antigen presentation. *Proc Natl Acad Sci U S A* 112, E1116-1125 (2015).

30. 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).

31. Follicular Lymphoma - Cancer Stat Facts. *SEER* https://seer.cancer.gov/statfacts/html/follicular.html.

32. Cheung, K.-J. J. *et al.* Acquired TNFRSF14 mutations in follicular lymphoma are associated with worse prognosis. *Cancer Res* 70, 9166–9174 (2010).

33. Pangault, C. *et al.* Integrative Analysis of Cell Crosstalk within Follicular Lymphoma Cell Niche: Towards a Definition of the FL Supportive Synapse. *Cancers (Basel)* 12, (2020).

34. Brien, G. L. *et al.* Simultaneous disruption of PRC2 and enhancer function underlies histone H3.3-K27M oncogenic activity in human hindbrain neural stem cells. *Nature Genetics* 53, 1221–1232 (2021).

35. Piunti, A. *et al.* Therapeutic targeting of polycomb and BET bromodomain proteins in diffuse intrinsic pontine gliomas. *Nat Med* 23, 493–500 (2017).

36. Wiese, M. *et al.* No Significant Cytotoxic Effect of the EZH2 Inhibitor Tazemetostat (EPZ-6438) on Pediatric Glioma Cells with Wildtype Histone 3 or Mutated Histone 3.3. *Klin Padiatr* 228, 113–117 (2016).

37. Westergaard Højfeldt, J. *et al.* Non-core subunits of the PRC2 complex are collectively required for its target site specificity. *Mol Cell* 76, 423-436.e3 (2019).

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

39. Knutson, S. K. *et al.* A selective inhibitor of EZH2 blocks H3K27 methylation and kills mutant lymphoma cells. *Nat Chem Biol* 8, 890–896 (2012).

40. Brach, D. *et al.* EZH2 Inhibition by Tazemetostat Results in Altered Dependency on B-cell Activation Signaling in DLBCL. *Mol Cancer Ther* 16, 2586–2597 (2017).

41. Sungalee, S. *et al.* Histone acetylation dynamics modulates chromatin conformation and allele-specific interactions at oncogenic loci. *Nature Genetics* 53, 650–662 (2021).

42. Ennishi, D. *et al.* Molecular and Genetic Characterization of MHC Deficiency Identifies EZH2 as Therapeutic Target for Enhancing Immune Recognition. *Cancer Discov* 9, 546–563 (2019).

43. Green, M. R. *et al.* Hierarchy in somatic mutations arising during genomic evolution and progression of follicular lymphoma. *Blood* 121, 1604–1611 (2013).

44. Jiang, Y. *et al.* CREBBP Inactivation Promotes the Development of HDAC3-Dependent Lymphomas. *Cancer Discov* 7, 38–53 (2017).

45. Wojcik, J. B. *et al.* Epigenomic Reordering Induced by Polycomb Loss Drives Oncogenesis but Leads to Therapeutic Vulnerabilities in Malignant Peripheral Nerve Sheath Tumors. *Cancer Res* 79, 3205–3219 (2019).

Figures, Supplemental data & legends

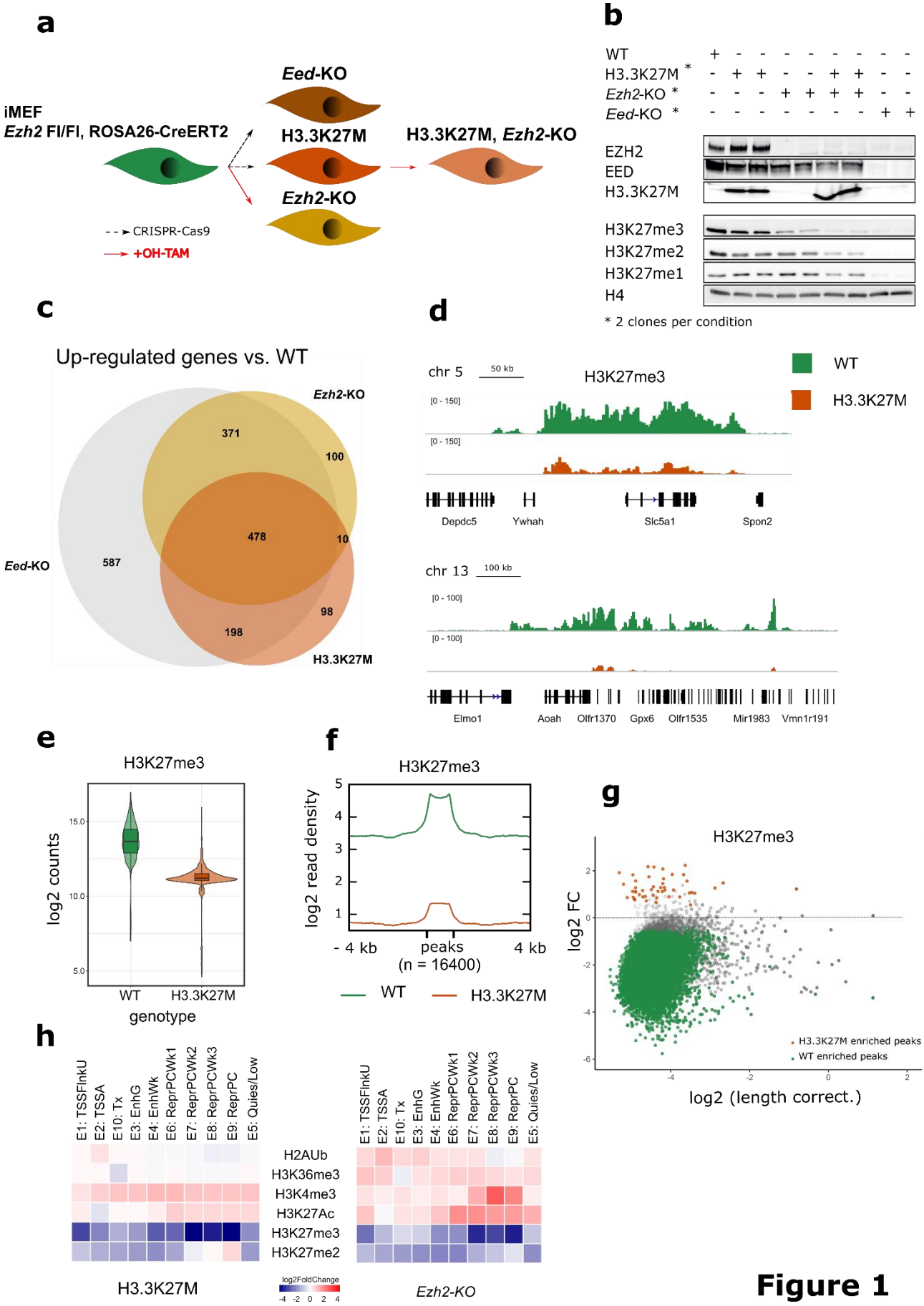

a
iMEF
Ezh2 Fl/Fl, ROSA26-CreERT2
Eed-KO
H3.3K27M
Ezh2-KO
H3.3K27M, Ezh2-KO
CRISPR-Cas9
+OH-TAM
b
WT
H3.3K27M *
Ezh2-KO *
Eed-KO *
EZH2
EED
H3.3K27M
H3K27me3
H3K27me2
H3K27me1
H4
* 2 clones per condition
c
Up-regulated genes vs. WT
Ezh2-KO
Eed-KO
H3.3K27M
371
100
478
10
587
98
198
d
chr 5
50 kb
H3K27me3
WT
H3.3K27M
Depdc5
Ywhah
Slc5a1
Spon2
chr 13
100 kb
Elmo1
Aoah
Olfr1370
Gpx6
Olfr1535
Mir1983
Vmn1r191
e
H3K27me3
log2 counts
WT
H3.3K27M
genotype
f
H3K27me3
log2 read density
- 4 kb
peaks
(n = 16400)
4 kb
WT
H3.3K27M
g
H3K27me3
log2 FC
H3.3K27M enriched peaks
WT enriched peaks
log2 (length correct.)
h
E1: TSSFlnkU
E2: TSSA
E10: Tx
E3: EnhG
E4: EnhWk
E6: ReprPCWk1
E7: ReprPCWk2
E8: ReprPCWk3
E9: ReprPC
E5: Quies/Low
H2AUb
H3K36me3
H3K4me3
H3K27Ac
H3K27me3
H3K27me2
H3.3K27M
log2FoldChange
Ezh2-KO

Figure 1

Figure 1

(a) Introduction in iMEF*Ezh2*fl/fl, ROSA26-CreERT2 of a stop cassette enabling the generation of *Eed*-KO cells, of H3.3K27M cells (point mutation) and a conditional deletion of *Ezh2* upon adjunction of 4-hydroxy tamoxifen (OH-TAM).

(b) Western blot analysis of EZH2, EED, H3.3K27M, H3K27me3, H3K27me2and H3K27me1 in WT, H3.3K27M cells with/without treatment with TAM (i.e. H3.3K27M; *Ezh2*-KO), *Ezh2*-KO and *Eed*-KO cells. H4 is used as a loading control.

(c) Venn Diagram showing the overlap of differentially upregulated genes vs. WT with DESeq 2 analysis (log2FC >1, p<.05) in: *Eed*-KO, *Ezh2*-KO and H3.3K27M cells.

(d) Representative sampling of CUT&RUN genomic tracks showing H3K27me3 distribution in WT and H3.3K27M cells. Two replicates are merged per track.

(e) Coupled violin/box plots showing the log2FC value of read counts per 2MB-bin retrieved upon H3K27me3 CUT&RUN-Seq in WT and H3.3K27M cells. Represented plots are average value of 2 replicates per condition.

(f) Density plot of the log2FC value of H3K27me3 read counts per peak, scaled to an equivalent 1kb, including 4kb upstream/downstream, across the sum of the peaks retrieved in both WT and H3.3K237M cells (MACS2 algorithm, qv0.0001), after merge of 2 replicates per condition.

(g) MA-plot showing the relative enrichment of H3K27me3 in peaks in both WT and H3.3K237M cells (y axis, WT: log2FC<0, H3.3K27M: log2FC>0). Base line level of H3K27me3 in WT normalized on the length of the peak is represented on the x axis. Green dots represent peaks specifically enriched in the WT over the mutant condition, orange dots represent the peaks specifically enriched in the mutant over the WT condition.

(h) Heatmaps summarizing ChromHMM model of 10-state emission parameters of concatenated sets of WT and either H3.3K27M (left panel) or *Ezh2*-KO (right panel) cells. H2Aub, H3K36me3, H3K4me3, H3K27ac, H3K27me3, H3K27me2 CUT&RUN-Seq were used as input files. Each row corresponds to the relative enrichment level of a given mark compared to WT, each column corresponds to an emission state. The detail of each emission state is provided in ***Fig S1g***.

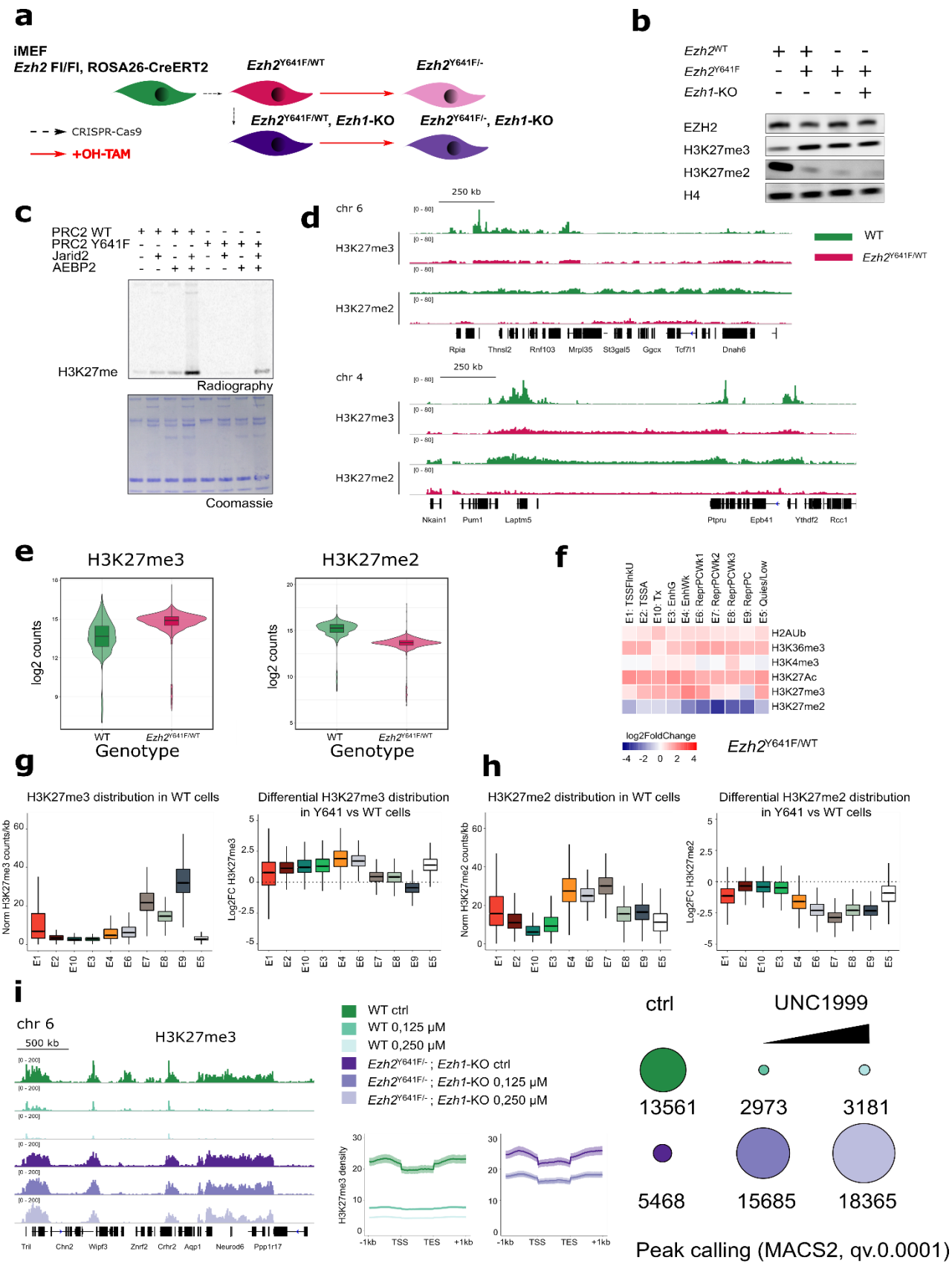

a
iMEF
Ezh2 FlFl, ROSA26-CreERT2
Ezh2Y641F/WT
Ezh2Y641F/-
Ezh2Y641F/WT, Ezh1-KO
Ezh2Y641F/-, Ezh1-KO
CRISPR-Cas9
+OH-TAM
b
Ezh2WT
Ezh2Y641F
Ezh1-KO
EZH2
H3K27me3
H3K27me2
H4
c
PRC2 WT
PRC2 Y641F
Jarid2
AEBP2
H3K27me
Radiography
Coomassie
d
chr 6
250 kb
H3K27me3
H3K27me2
WT
Ezh2Y641F/WT
Rpia
Thnsl2
Rnf103
Mrpl35
St3gal5
Ggcx
Tcf7l1
Dnah6
chr 4
Nkain1
Pum1
Laptm5
Ptpru
Epb41
Ythdf2
Rcc1
e
H3K27me3
H3K27me2
log2 counts
Genotype
f
E1: TSSFlnkU
E2: TSSA
E10: Tx
E3: EnhG
E4: EnhWk
E6: ReprPCWk1
E7: ReprPCWk2
E8: ReprPCWk3
E9: ReprPC
E5: Quies/Low
H2AUb
H3K36me3
H3K4me3
H3K27Ac
H3K27me3
H3K27me2
log2FoldChange
g
H3K27me3 distribution in WT cells
Differential H3K27me3 distribution in Y641 vs WT cells
Norm H3K27me3 counts/kb
Log2FC H3K27me3
h
H3K27me2 distribution in WT cells
Differential H3K27me2 distribution in Y641 vs WT cells
Norm H3K27me2 counts/kb
Log2FC H3K27me2
i
chr 6
500 kb
H3K27me3
Tril
Chn2
Wipf3
Znrf2
Chrr2
Aqp1
Neurod6
Ppp1r17
WT ctrl
WT 0,125 µM
WT 0,250 µM
Ezh2Y641F/- ; Ezh1-KO ctrl
Ezh2Y641F/- ; Ezh1-KO 0,125 µM
Ezh2Y641F/- ; Ezh1-KO 0,250 µM
H3K27me3 density
-1kb
TSS
TES
+1kb
ctrl
UNC1999
13561
2973
3181
5468
15685
18365
Peak calling (MACS2, qv.0.0001)

Figure 2

Figure 2

(a) Insertion in iMEF*Ezh2*fl/fl, ROSA26-CreERT2 of *Ezh2*^{Y641F} on one allele using CRISPR/Cas9 technology (cf ***Fig 2Sa***), further enabling the generation of *Ezh1*-KO cells and a conditional deletion of *Ezh2* WT allele upon adjunction of 4-hydroxy tamoxifen (OH-TAM).

(b) Western blot analysis of EZH2, H3K27me3 and H3K27me2 in WT, *Ezh2*$^{Y641F/WT}$, *Ezh2*$^{Y641F/-}$ and *Ezh2*$^{Y641F/-}$ *; Ezh*1-KO cells. H4 is used as a loading control.

(c) HMT assay performed with PRC2-EZH2 reconstituted with either WT or Y641F mutant *EZH2*, in the presence of peptides mimicking either JARID2 or AEBP1 or both.

(d) Representative sampling of CUT&RUN genomic tracks showing H3K27me3 and H3K27me2 deposition in WT and *Ezh2*$^{Y641F/WT}$ cells. Two replicates are merged per track.

(e) Coupled violin/box plots showing the log2FC value of read counts per 2MB-bin retrieved upon H3K27me3 (left panel) and H3K27me2 (right panel) CUT&RUN-Seq in WT and *Ezh2*$^{Y641F/WT}$ cells. Represented plots are average value of 2 replicates per condition.

(f) Heatmaps summarizing ChromHMM model of 10-state emission parameters of concatenated sets of WT and *Ezh2*$^{Y641F/WT}$ cells. Representation is similar to ***Fig 1h***.

(g) Box plot showing H3K27me3 signal in WT cells across chromatin states as defined in ***Fig S1g*** (left panel)***.*** Box plot showing differential H3K27me3 signal between *Ezh2*$^{Y641F/WT}$ cells (>log2FC) and WT cells (<log2FC) (right panel).

(h) Similar to ***Fig 2g***, with H3K27me2.

(i) Left panel: Representative sampling of CUT&RUN genomic tracks showing H3K27me3 enrichment in WT and *Ezh2*$^{Y641F/-}$ *; Ezh1*-KO cells upon increasing doses of UNC 1999. Middle panel: Density plot of the log2FC value of H3K27me3 read counts centered around all the genes, scaled to an equivalent 1kb, including 1kb upstream/downstream, throughout progressive inhibition of EZH2 in both WT and *Ezh2*$^{Y641F/-}$*Ezh1*-KO cells (MACS2 algorithm, qv0.0001). Plot of H3K27me3 in *Ezh2*$^{Y641F/-}$ *; Ezh1*-KO cells treated with 0.250 µM is encompassed within *Ezh2*$^{Y641F/-}$ *; Ezh1*-KO cells treated with 0.125 µM plot. Right panel: Number of peaks retrieved using peaks calling algorithm MACS2 (qv 0.0001) per condition, using the sum of the peaks from 2 replicates per condition.

Two replicates are merged per condition.

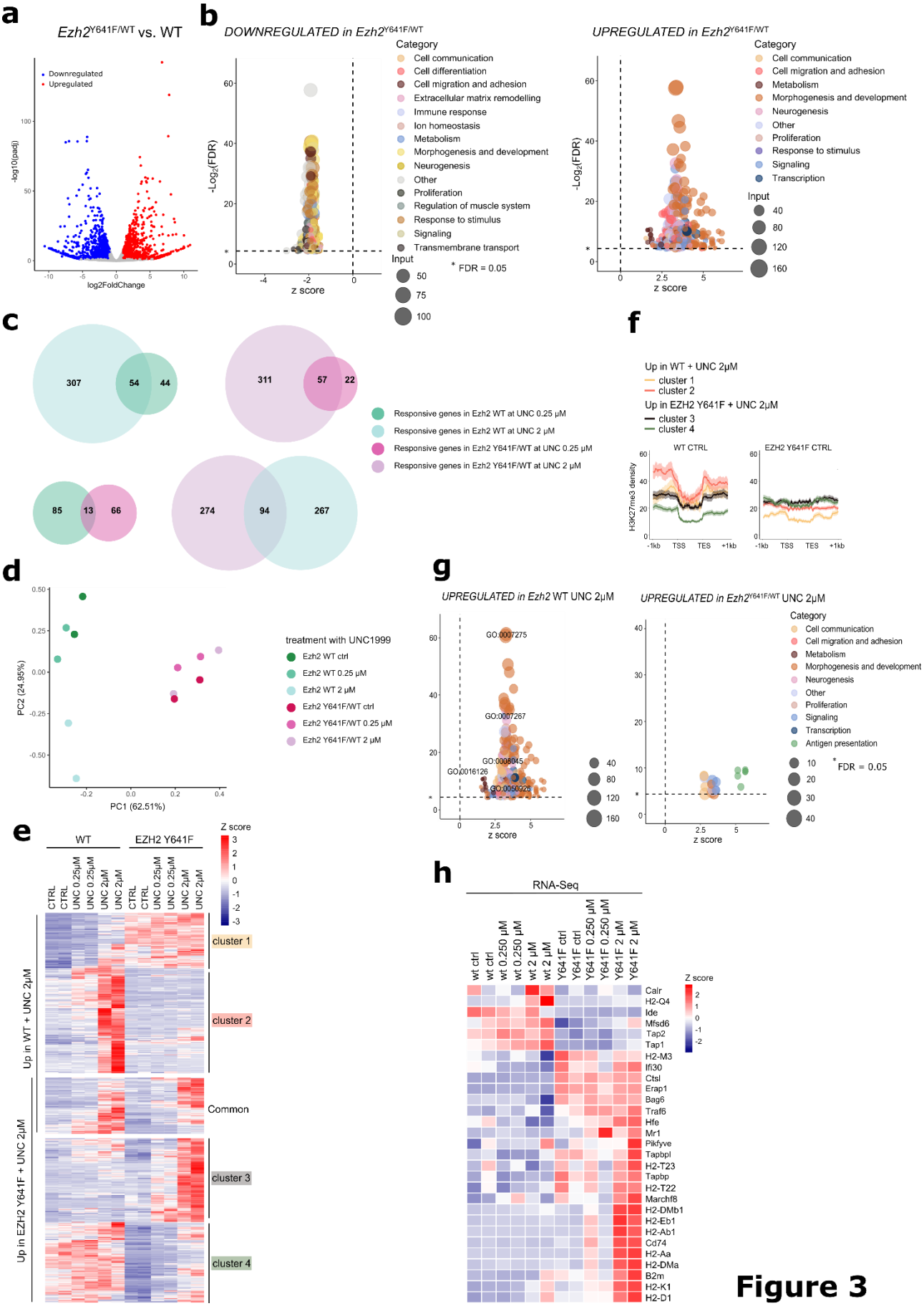

a
Ezh2Y641F/WT vs. WT
Downregulated
Upregulated
-log10(padj)
log2FoldChange
b
DOWNREGULATED in Ezh2Y641F/WT
UPREGULATED in Ezh2Y641F/WT
-Log2(FDR)
z score
Category
Cell communication
Cell differentiation
Cell migration and adhesion
Extracellular matrix remodelling
Immune response
Ion homeostasis
Metabolism
Morphogenesis and development
Neurogenesis
Other
Proliferation
Regulation of muscle system
Response to stimulus
Signaling
Transmembrane transport
Transcription
Input
* FDR = 0.05
c
307
54
44
311
57
22
85
13
66
274
94
267
Responsive genes in Ezh2 WT at UNC 0.25 µM
Responsive genes in Ezh2 WT at UNC 2 µM
Responsive genes in Ezh2 Y641F/WT at UNC 0.25 µM
Responsive genes in Ezh2 Y641F/WT at UNC 2 µM
f
Up in WT + UNC 2µM
cluster 1
cluster 2
Up in EZH2 Y641F + UNC 2µM
cluster 3
cluster 4
WT CTRL
EZH2 Y641F CTRL
H3K27me3 density
-1kb
TSS
TES
+1kb
d
PC2 (24.95%)
PC1 (62.51%)
treatment with UNC1999
Ezh2 WT ctrl
Ezh2 WT 0.25 µM
Ezh2 WT 2 µM
Ezh2 Y641F/WT ctrl
Ezh2 Y641F/WT 0.25 µM
Ezh2 Y641F/WT 2 µM
g
UPREGULATED in Ezh2 WT UNC 2µM
UPREGULATED in Ezh2Y641F/WT UNC 2µM
GO:0007275
GO:0007267
GO:0008045
GO:0016126
GO:0050926
z score
Category
Cell communication
Cell migration and adhesion
Metabolism
Morphogenesis and development
Neurogenesis
Other
Proliferation
Signaling
Transcription
Antigen presentation
* FDR = 0.05
e
WT
EZH2 Y641F
CTRL
UNC 0.25µM
UNC 2µM
Z score
cluster 1
cluster 2
Common
cluster 3
cluster 4
Up in WT + UNC 2µM
Up in EZH2 Y641F + UNC 2µM
h
RNA-Seq
wt ctrl
wt 0.250 µM
wt 2 µM
Y641F ctrl
Y641F 0.250 µM
Y641F 2 µM
Z score
Calr
H2-Q4
Ide
Mfsd6
Tap2
Tap1
H2-M3
Ifi30
Ctsl
Erap1
Bag6
Traf6
Hfe
Mr1
Pikfyve
Tapbpl
H2-T23
Tapbp
H2-T22
Marchf8
H2-DMb1
H2-Eb1
H2-Ab1
Cd74
H2-Aa
H2-DMa
B2m
H2-K1
H2-D1

Figure 3

Figure 3

(a) Volcano plot showing the DEGs in *Ezh2*$^{Y641F/WT}$ vs. WT cells, after merge of two replicates per condition. Red dots represent upregulated genes and blue dots represent downregulated genes.

(b) Bubble plots showing GO terms enriched in downregulated (left panel) and upregulated (right panel) gene sets in *Ezh2*$^{Y641F/WT}$ cells compared to WT. Input size corresponds to the number of genes from the gene set represented in a given broad category (see ***Material & methods***).

(c) Venn Diagrams showing overlap of differentially upregulated genes (i.e. responsive genes) upon UNC199 in *Ezh2*$^{Y641F/WT}$ and WT cells, vs. control. Upper panel displays the comparison of gene sets according to the level of inhibition of EZH2 within one same genotype, lower panel displays the comparison of gene sets according to the genotype for a given level of inhibition of EZH2.

(d) Principal Component Analysis showing the reparation of the RNA-Seq data from both *Ezh2*$^{Y641F/WT}$ and WT cells, either non-treated (ctrl), or treated with increasing doses of UNC1999.

(e) Heatmap of RNA-Seq expression z-scores computed for DEG in *Ezh2*$^{Y641F/WT}$ vs. WT cells either non-treated (ctrl), or treated with increasing doses of UNC1999. Upper panel (clusters 1 & 2) corresponds to genes that are responsive to EZH2 inhibition in WT cells and shows their expression level in the mutant counterpart. Middle panel shows the level of expression of genes responsive in both genotypes (Common). Lower panel (clusters 3 & 4) corresponds to genes that are responsive to EZH2 inhibition in mutant cells and shows their expression level in the WT cells.

(f) Density plots of the log2FC value of H3K27me3 read counts centered around the gene, scaled to an equivalent 1kb, including 1kb upstream/downstream, across the list of responsive genes identified in 4 clusters, at baseline level (prior to inhibition), in WT (left panel) and *Ezh2*$^{Y641F/WT}$ cells (right panel).

(g) Bubble plots showing GO terms enriched in upregulated/responsive genes in WT cells treated with UNC1999 2µM (left panel) and upregulated/responsive genes in *Ezh2*$^{Y641F/WT}$ cells treated with UNC1999 2µM (right panel). Mode of representation is similar to ***Fig3b***. Extensive data set of this analysis are available in ***FigS3b-d***.

(h) Heatmap of RNA-Seq expression z-scores computed for genes from the most enriched GO term in *Ezh2*$^{Y641F/WT}$ cells treated with UNC 1999 2µM, GO Term GO:0019886 (antigen processing and presentation of exogenous peptide antigen *via* MHC class II).

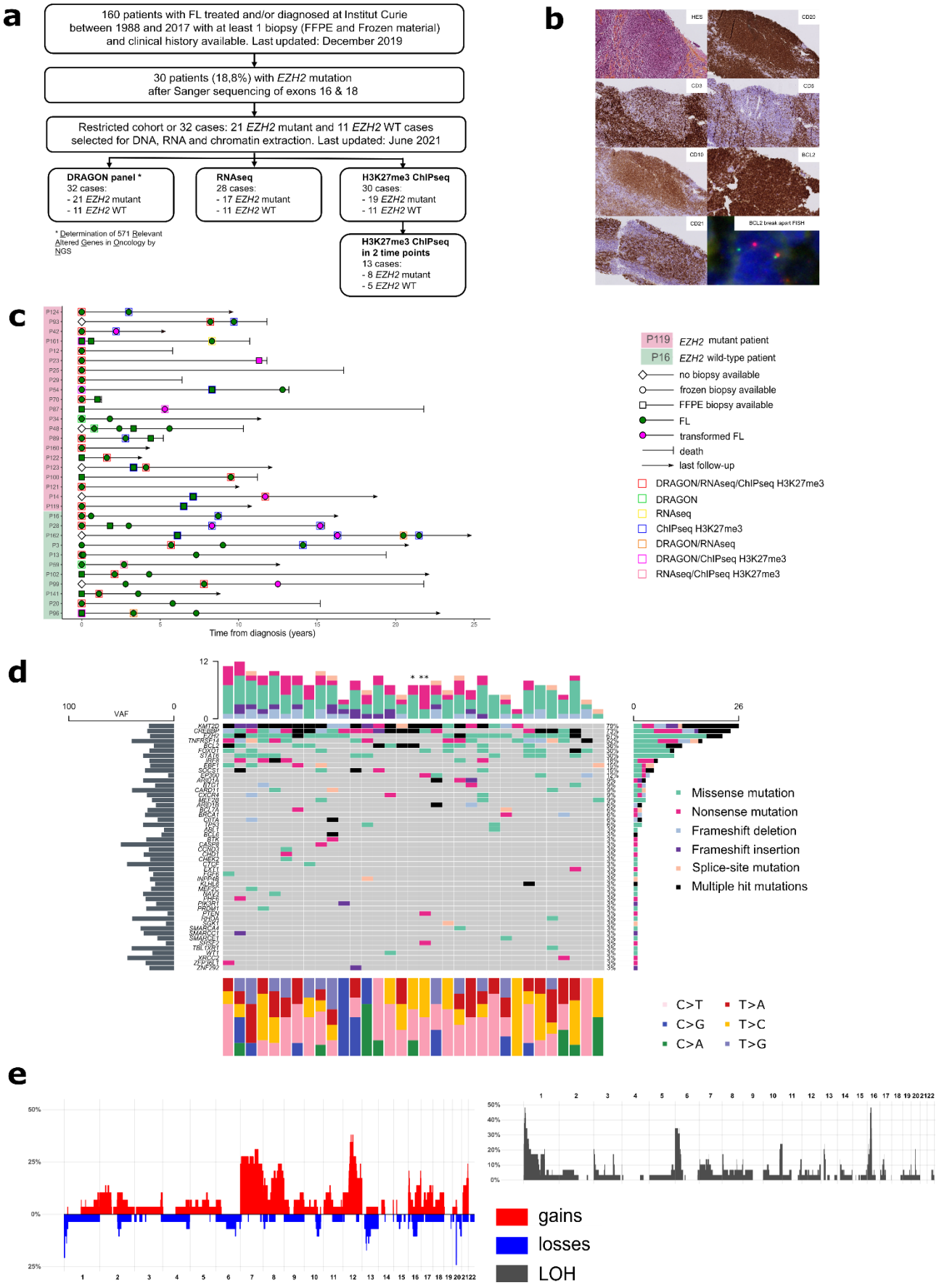

a
160 patients with FL treated and/or diagnosed at Institut Curie between 1988 and 2017 with at least 1 biopsy (FFPE and Frozen material) and clinical history available. Last updated: December 2019
30 patients (18,8%) with EZH2 mutation after Sanger sequencing of exons 16 & 18
Restricted cohort or 32 cases: 21 EZH2 mutant and 11 EZH2 WT cases selected for DNA, RNA and chromatin extraction. Last updated: June 2021
DRAGON panel *
32 cases:
- 21 EZH2 mutant
- 11 EZH2 WT
RNAseq
28 cases:
- 17 EZH2 mutant
- 11 EZH2 WT
H3K27me3 ChIPseq
30 cases:
- 19 EZH2 mutant
- 11 EZH2 WT
* Determination of 571 Relevant Altered Genes in Oncology by NGS
H3K27me3 ChIPseq in 2 time points
13 cases:
- 8 EZH2 mutant
- 5 EZH2 WT
b
HES
CD20
CD3
CD5
CD10
BCL2
CD21
BCL2 break apart FISH
c
Time from diagnosis (years)
P119 EZH2 mutant patient
P16 EZH2 wild-type patient
no biopsy available
frozen biopsy available
FFPE biopsy available
FL
transformed FL
death
last follow-up
DRAGON/RNAseq/ChIPseq H3K27me3
DRAGON
RNAseq
ChIPseq H3K27me3
DRAGON/RNAseq
DRAGON/ChIPseq H3K27me3
RNAseq/ChIPseq H3K27me3
d
VAF
Missense mutation
Nonsense mutation
Frameshift deletion
Frameshift insertion
Splice-site mutation
Multiple hit mutations
C>T
C>G
C>A
T>A
T>C
T>G
e
gains
losses
LOH

Figure 4

Figure 4

(a) Workflow of cohort sample selection and downstream analyses.

(b) Histological (Hematoxylin Eosin Saffron, X10), immunohistochemical (CD20, CD3, CD5, CD10, BCL2, CD21) and cytogenetic (BCL2 break-apart probe, Fluorescence In Situ, Hybridization) features of follicular lymphoma. Example in one lymph node core needle biopsy from P#154, reagents & imaging: department of Pathology, Institut Curie).

(c) Biopsy information with disease event timelines for the restricted cohort of 32 cases.

(d) Frequency and type of mutations affecting 48 genes identified throughout the restricted cohort of 32 cases. Each column represents one patient, each row one gene. * and ** columns correspond to samples from patient P#96, sequenced at the time of diagnosis and time of first relapse respectively. Bar graph on the top represents the number of mutations per patient, bar graph on the left represents mean VAF per gene, bar graph on the right represents the number and type of mutations per gene, bar graph on the bottom represents the distribution of base substitution patterns for all pathologic variants identified per patient. The mutation waterfall plot was created using the maftools package in R.

(e) Summary of CNV frequency and type (left panel: gains and losses, right panel: loss of heterozygosity) throughout the restricted cohort of 32 cases. Percentages of cases affected are represented on the left of the plots.

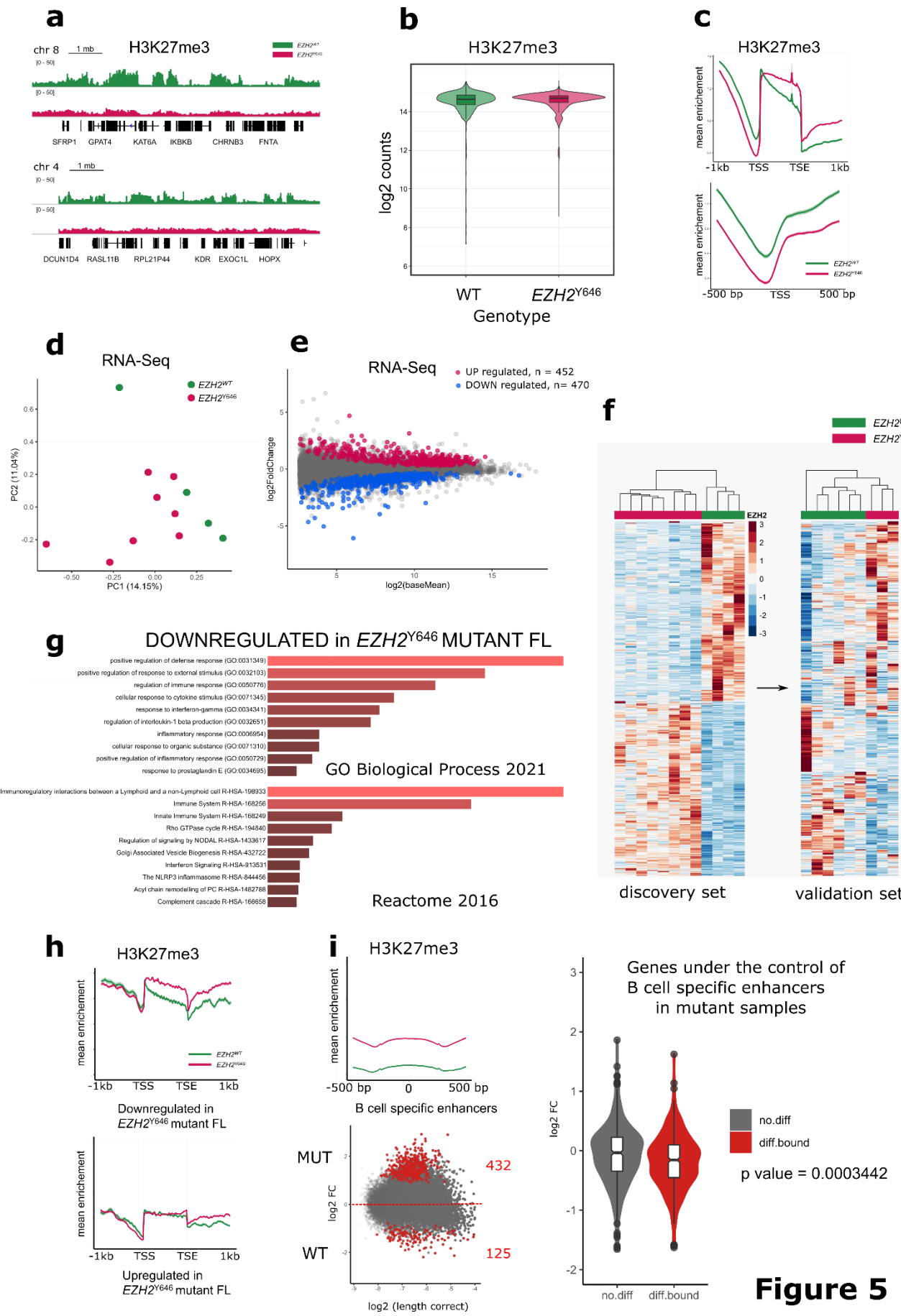

a
H3K27me3
chr 8
1 mb
SFRP1 GPAT4 KAT6A IKBKB CHRNB3 FNTA
chr 4
1 mb
DCUN1D4 RASL11B RPL21P44 KDR EXOC1L HOPX
b
H3K27me3
log2 counts
WT
EZH2Y646
Genotype
c
H3K27me3
mean enrichement
-1kb TSS TSE 1kb
-500 bp TSS 500 bp
d
RNA-Seq
EZH2WT
EZH2Y646
PC2 (11.04%)
PC1 (14.15%)
e
RNA-Seq
UP regulated, n = 452
DOWN regulated, n= 470
log2FoldChange
log2(baseMean)
f
EZH2WT
EZH2Y646
EZH2
discovery set
validation set
g
DOWNREGULATED in EZH2Y646 MUTANT FL
positive regulation of defense response (GO:0031349)
positive regulation of response to external stimulus (GO:0032103)
regulation of immune response (GO:0050776)
cellular response to cytokine stimulus (GO:0071345)
response to interferon-gamma (GO:0034341)
regulation of interleukin-1 beta production (GO:0032651)
inflammatory response (GO:0006954)
cellular response to organic substance (GO:0071310)
positive regulation of inflammatory response (GO:0050729)
response to prostaglandin E (GO:0034695)
GO Biological Process 2021
Immunoregulatory interactions between a Lymphoid and a non-Lymphoid cell R-HSA-198933
Immune System R-HSA-168256
Innate Immune System R-HSA-168249
Rho GTPase cycle R-HSA-194840
Regulation of signaling by NODAL R-HSA-1433617
Golgi Associated Vesicle Biogenesis R-HSA-432722
Interferon Signaling R-HSA-913531
The NLRP3 inflammasome R-HSA-844456
Acyl chain remodelling of PC R-HSA-1482788
Complement cascade R-HSA-166658
Reactome 2016
h
H3K27me3
mean enrichement
-1kb TSS TSE 1kb
Downregulated in
EZH2Y646 mutant FL
Upregulated in
EZH2Y646 mutant FL
i
H3K27me3
-500 bp 0 500 bp
B cell specific enhancers
MUT
WT
432
125
log2 FC
log2 (length correct)
Genes under the control of
B cell specific enhancers
in mutant samples
log2 FC
no.diff
diff.bound
p value = 0.0003442

Figure 5

Figure 5

All the analyses were conducted in 4 *EZH2* WT and 8 *EZH2*^{Y6416} FL samples, after merge of 2 replicates per patient for ChIP-Seq experiments.

(a) Representative sampling of ChIP-Seq genomic tracks showing H3K27me3 deposition in one WT and one *EZH2*^{Y6416} tumor sample.

(b) Coupled violin/box plots showing the log2FC value of read counts per 2MB-bin, RPKM normalized, retrieved upon H3K27me3 ChIP-Seq in WT and *EZH2*^{Y646} patient samples.

(c) Upper panel: Density plot of the log2FC value of H3K27me3 read counts centered around the gene body scaled to an equivalent 1kb, including 1kb upstream/downstream. Lower panel: Density plot of the log2FC value of H3K27me3 read counts centered around the TSS, including 500 bp upstream/downstream.

(d) Principal Component Analysis showing the repartition of the RNA-Seq data from both *EZH2*^{Y6416} and WT FL samples.

(e) MA-plot showing differentially expressed genes between *EZH2* WT and *EZH2*^{Y6416} FL samples. Significantly upregulated genes in mutant vs. WT cases are represented by red dots, significantly downregulated genes in mutant vs. WT cases are represented by blue dots.

(f) Left panel: heatmap of hierarchical clustering of *EZH2* WT and *EZH2*^{Y6416} FL samples with DEG from the discovery set of 12 patients. Right panel: This signature properly segregates *EZH2* WT and *EZH2*^{Y6416} cases in a validation set of 9 cases.

(g) Upper panel:10 most significantly enriched GO terms upon downregulated genes in mutant cases vs. WT (GO biological processes, Enrichr/Panther 15.0). Lower panel: 10 most significantly enriched pathways upon downregulated genes in mutant cases vs WT (Enrichr/Reactome database release 77).

(h) Density-plot of the log2FC value of H3K27me3 read counts centered around the gene body of downregulated (upper panel) and upregulated (lower panel) genes scaled to an equivalent 1kb, including 1kb upstream/downstream.

(i) Upper left panel: Density-plot of the log2FC value of H3K27me3 read counts centered around B cell specific enhancer (defined as the intersection of .bed files for H3K27ac and H3K4me1 enrichment in B cells available on https://www.blueprint-epigenome.eu, 2 replicates per condition), including 500 bp upstream/downstream. Lower left panel: MA-plot showing the enrichment in H3K27me3 according to the genotype (y axis, WT: log2FC<0, *EZH2*^{Y646}: log2FC>0) at B cell specific enhancers. Baseline enrichment in H3K27me3 in WT samples normalized on the length of the enhancer serves as a reference (x axis, log2 read counts/length of the enhancer). Significantly differentially bound enhancers are represented by red dots (padj<0.05). Right panel: Coupled violin/box plots showing the log2FC value of RNA-Seq read counts of genes related to B cell specific enhancers either differentially bound with in H3K27me3 in presence of the mutation (red plot), or non-differentially bound in presence of the mutation (grey plot -0.5<log2FC<0.5, random subsampling of 432 enhancers).

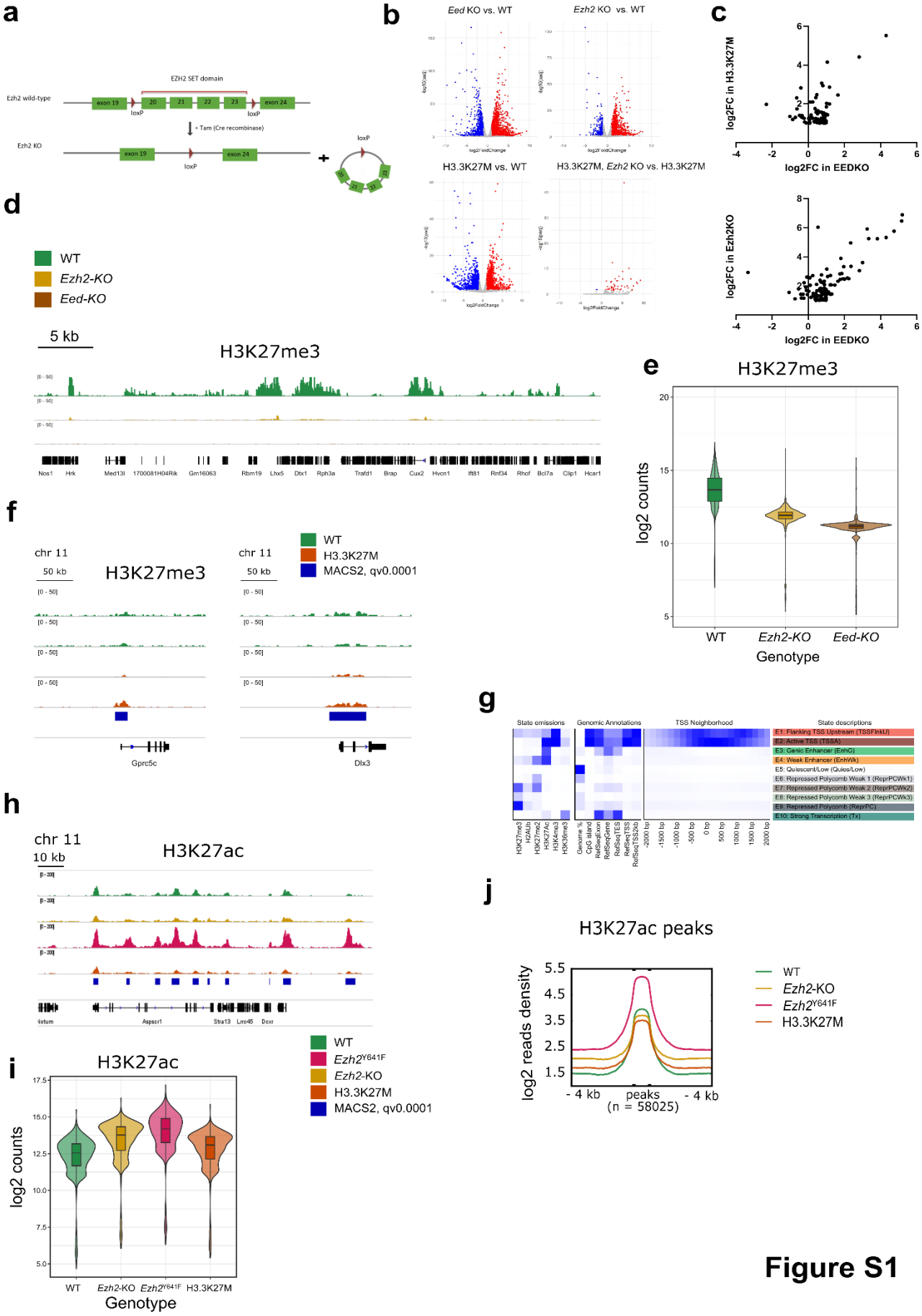

a
EZH2 SET domain
Ezh2 wild-type
exon 19
20
21
22
23
exon 24
loxP
loxP
+ Tam (Cre recombinase)
Ezh2 KO
exon 19
exon 24
loxP
loxP
b
Eed KO vs. WT
Ezh2 KO vs. WT
H3.3K27M vs. WT
H3.3K27M, Ezh2 KO vs. H3.3K27M
c
log2FC in H3.3K27M
log2FC in EEDKO
log2FC in Ezh2KO
log2FC in EEDKO
d
WT
Ezh2-KO
Eed-KO
5 kb
H3K27me3
Nos1
Hrk
Med13l
1700081H04Rik
Gm16063
Rbm19
Lhx5
Dtx1
Rph3a
Trafd1
Brap
Cux2
Hvcn1
Ift81
Rnf34
Rhof
Bcl7a
Clip1
Hcar1
e
H3K27me3
log2 counts
WT
Ezh2-KO
Eed-KO
Genotype
f
chr 11
50 kb
H3K27me3
chr 11
50 kb
WT
H3.3K27M
MACS2, qv0.0001
Gprc5c
Dlx3
g
State emissions
Genomic Annotations
TSS Neighborhood
State descriptions
E1: Flanking TSS Upstream (TSSFlnkU)
E2: Active TSS (TSSA)
E3: Genic Enhancer (EnhG)
E4: Weak Enhancer (EnhWk)
E5: Quiescent/Low (Quies/Low)
E6: Repressed Polycomb Weak 1 (ReprPCWk1)
E7: Repressed Polycomb Weak 2 (ReprPCWk2)
E8: Repressed Polycomb Weak 3 (ReprPCWk3)
E9: Repressed Polycomb (ReprPC)
E10: Strong Transcription (Tx)
h
chr 11
10 kb
H3K27ac
Vetum
Aspscr1
Stra13
Lrrc45
Dexr
WT
Ezh2Y641F
Ezh2-KO
H3.3K27M
MACS2, qv0.0001
i
H3K27ac
log2 counts
WT
Ezh2-KO
Ezh2Y641F
H3.3K27M
Genotype
j
H3K27ac peaks
log2 reads density
- 4 kb
peaks
(n = 58025)
- 4 kb
WT
Ezh2-KO
Ezh2Y641F
H3.3K27M

Figure S1

Figure S1

(a) Schematic strategy of *Ezh2* SET domain-conditional allele deletion and cre-mediated recombination upon OH-TAM treatment.

(b) Volcano plots showing the DEGs in *Eed*-KO vs. WT cells (top left panel), *Ezh2*-KO vs. WT cells (top right panel), H3.3K27M vs. WT cells (bottom left panel), H3.3K27M vs. H3.3K27M *Ezh2*-KO cells (bottom right panel). Red dots represent upregulated genes and blue dots represent downregulated genes in the formerly mentioned condition.

(c) Correlation plots of the expression level vs WT (log2FC) of 108 genes significantly upregulated in H3.3K27M and non-overlapping with *Eed*-KO upregulated genes (upper graph) and of 110 genes significantly upregulated in *Ezh2*-KO and non-overlapping with *Eed*-KO upregulated genes (lower graph).

(d) Representative sampling of CUT&RUN genomic tracks showing H3K27me3 deposition in WT, *Ezh2*-KO and *Eed*-KO cells. Two replicates are merged per track.

(e) Coupled violin/box plots showing the log2FC value of read counts per 2MB-bin retrieved upon H3K27me3 CUT&RUN-Seq in WT, *Ezh2*-KO and *Eed*-KO cells. Represented plots are average value of 2 replicates per condition.

(f) Sampling of CUT&RUN genomic tracks showing H3K27me3 deposition in WT and H3.3K27M cells, at peaks significantly enriched in H3.3K27M cells over WT. Two replicates per condition are shown separately, defined peaks are represented by blue bars (bottom track).

(g) Heatmaps summarizing ChromHMM model of 10-state emission parameters defined according to level of enrichment of various marks (H2Aub, H3K36me3, H3K4me3, H3K27ac, H3K27me3, H3K27me2) in regard of annotated regions. Active states are E1, E2 and E10. Enhancers are subdivided in genic enhancers (E3) and weak enhancer (E4). Strongly repressed state E9 correspond to canonical Polycomb targets, mildest repression corresponds to Polycomb weak states (E6, E7, E8). Quiescent state is E10.

(h) Representative sampling of CUT&RUN genomic tracks showing H3K27ac deposition in WT, *Ezh2*-KO, *Ezh2*$^{Y641F/WT}$ and H3.3K27M cells. Two replicates are merged per track. MACS2-defined H3K27ac peaks as the sum of the peaks detected in the 4 conditions are represented by blue bars (bottom track).

(i) Coupled violin/box plots showing the log2FC value of read counts per 2MB-bin retrieved upon H3K27ac CUT&RUN-Seq in WT, *Ezh2*-KO, *Ezh2*$^{Y641F/WT}$ and H3.3K27M cells. Represented plots are average value of 2 replicates per condition.

(j) Density plots of the log2FC value of H3K27ac read count per peaks, scaled to an equivalent 1kb, including 4kb upstream/downstream, across the sum of the peaks retrieved in WT, *Ezh2*-KO, *Ezh2*$^{Y641F/WT}$ and H3.3K27M cells (MACS2 algorithm, qv0.0001), after merge of 2 replicates per condition.

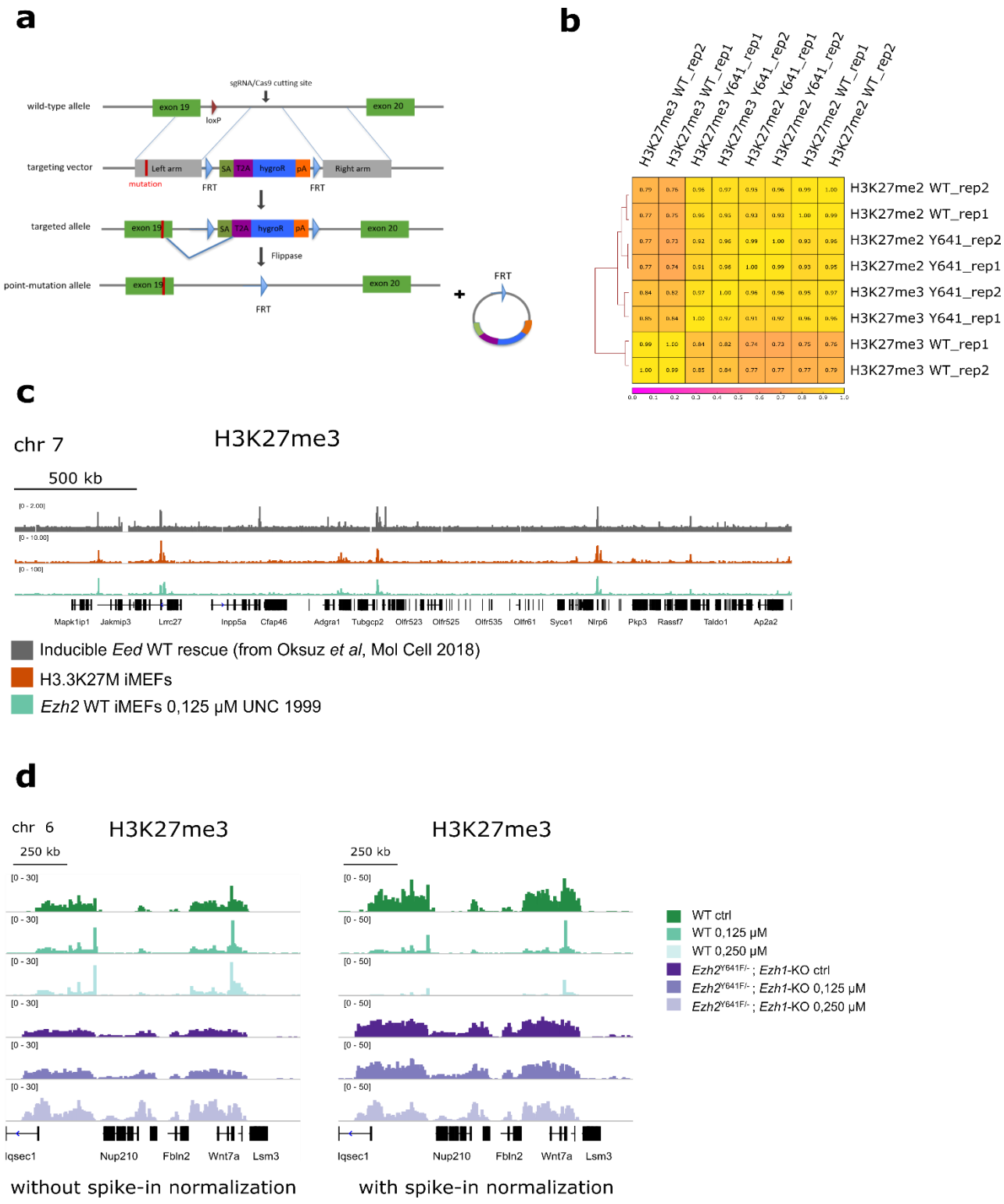

a
sgRNA/Cas9 cutting site
wild-type allele
exon 19
loxP
exon 20
targeting vector
Left arm
SA
T2A
hygroR
pA
Right arm
mutation
FRT
FRT
targeted allele
Flippase
point-mutation allele
FRT
FRT
b
H3K27me3 WT_rep2
H3K27me3 WT_rep1
H3K27me3 Y641_rep1
H3K27me3 Y641_rep2
H3K27me2 Y641_rep1
H3K27me2 Y641_rep2
H3K27me2 WT_rep1
H3K27me2 WT_rep2
H3K27me2 WT_rep2
H3K27me2 WT_rep1
H3K27me2 Y641_rep2
H3K27me2 Y641_rep1
H3K27me3 Y641_rep2
H3K27me3 Y641_rep1
H3K27me3 WT_rep1
H3K27me3 WT_rep2
c
chr 7
H3K27me3
500 kb
Mapk1ip1
Jakmip3
Lrrc27
Inpp5a
Cfap46
Adgra1
Tubgcp2
Olfr523
Olfr525
Olfr535
Olfr61
Syce1
Nlrp6
Pkp3
Rassf7
Taldo1
Ap2a2
Inducible Eed WT rescue (from Oksuz et al, Mol Cell 2018)
H3.3K27M iMEFs
Ezh2 WT iMEFs 0,125 µM UNC 1999
d
chr 6
H3K27me3
250 kb
Iqsec1
Nup210
Fbln2
Wnt7a
Lsm3
without spike-in normalization
H3K27me3
250 kb
with spike-in normalization
WT ctrl
WT 0,125 µM
WT 0,250 µM
Ezh2Y641F/- ; Ezh1-KO ctrl
Ezh2Y641F/- ; Ezh1-KO 0,125 µM
Ezh2Y641F/- ; Ezh1-KO 0,250 µM

Figure S2

Figure S2

(a) Schematic representation of the insertion of *Ezh2*^{Y641F} mutation by homologous recombination. In brief, the right arm partly overlaps with *Ezh2* exon 16 in which an adenine is replaced by a thymine, resulting in the mutation of the tyrosine Y641 in a phenylalanine. A selection cassette is flanked by FRT sequences, which allow its excision upon transfection with flippase (FRT-FLP recombination). SA: Splicing Acceptor; T2A: T2A sequence; HygroR: Hygromycin resistance; pA: polyA sequence.

(b) Correlation matrix of the replicates of H3K27me2 and H3K27me3 CUT&RUNN-Seq data from WT and *Ezh2*$^{Y641F/WT}$ cells, showing the proximity of H3K27me3 in mutant cells with H3K27me2 in WT cells.

(c) Representative sampling of CUT&RUN genomic tracks showing H3K27me3 deposition in mESC (C57BL/6, inducible WT EDD/rescue cells, from Oksuz *et al*, Mol Cell 2018, *Capturing the onset of PRC2-mediated repressive domain formation, one replicate*), H3.3K27M cells and *Ezh2*^{WT} cells treated with 0.125 µM of UNC1999. Two replicates are merged in the two bottom tracks.

(d) Representative sampling of CUT&RUN genomic tracks showing H3K27me3 deposition in WT and *Ezh2*$^{Y641F/-}$ *;Ezh1*-KO cells upon increasing doses of UNC1999 without normalization on *Drosophila* spikes (left panel) and with normalization (right panel). Two replicates are merged per track.

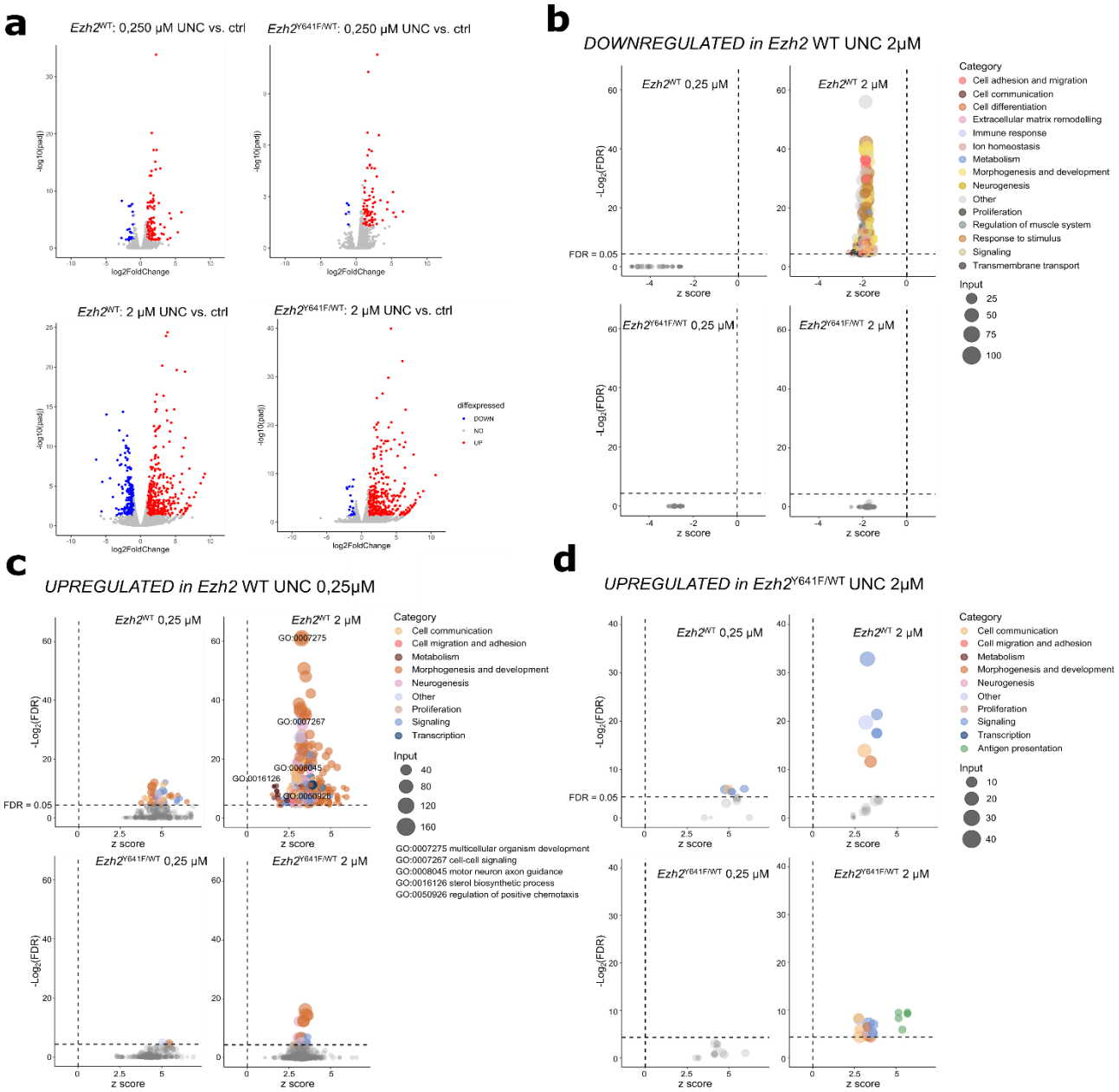

a
Ezh2WT: 0,250 µM UNC vs. ctrl
Ezh2Y641F/WT: 0,250 µM UNC vs. ctrl
Ezh2WT: 2 µM UNC vs. ctrl
Ezh2Y641F/WT: 2 µM UNC vs. ctrl
-log10(padj)
log2FoldChange
diffexpressed
DOWN
NO
UP
b
DOWNREGULATED in Ezh2 WT UNC 2µM
Ezh2WT 0,25 µM
Ezh2WT 2 µM
Ezh2Y641F/WT 0,25 µM
Ezh2Y641F/WT 2 µM
-Log2(FDR)
FDR = 0.05
z score
Category
Cell adhesion and migration
Cell communication
Cell differentiation
Extracellular matrix remodelling
Immune response
Ion homeostasis
Metabolism
Morphogenesis and development
Neurogenesis
Other
Proliferation
Regulation of muscle system
Response to stimulus
Signaling
Transmembrane transport
Input
25
50
75
100
c
UPREGULATED in Ezh2 WT UNC 0,25µM
GO:0007275
GO:0007267
GO:0008045
GO:0016126
GO:0050926
Cell communication
Cell migration and adhesion
Metabolism
Morphogenesis and development
Neurogenesis
Other
Proliferation
Signaling
Transcription
40
80
120
160
GO:0007275 multicellular organism development
GO:0007267 cell-cell signaling
GO:0008045 motor neuron axon guidance
GO:0016126 sterol biosynthetic process
GO:0050926 regulation of positive chemotaxis
d
UPREGULATED in Ezh2Y641F/WT UNC 2µM
Cell communication
Cell migration and adhesion
Metabolism
Morphogenesis and development
Neurogenesis
Other
Proliferation
Signaling
Transcription
Antigen presentation
10
20
30
40

Figure S3

Figure S3

(a) Volcano plots showing the DEGs in treated cells vs. ctrl cells: WT cells treated with UNC 0.25µM (top left panel) or 2µM (bottom left panel), *Ezh2*$^{Y641F/WT}$ cells treated with UNC 0.25µM (top right panel) or 2µM (bottom right panel). Red dots represent upregulated genes and blue dots represent downregulated genes.

(b) Extended data from ***Fig 3h***: bubble plots showing the level of enrichment across 4 sample types (in WT cells treated with either 0.25µM or 2 µM and in *Ezh2*$^{Y641F/WT}$ cells treated the same way) in GO terms found to be significantly enriched in downregulated gene set from WT cells treated with 2µM compared with ctrl. As an example, GO terms enriched in downregulated genes in WT cells treated with 2µM (top right panel) are not enriched in any of the 3 other conditions.

(c) & (d): Similar to ***Fig S3b,*** focusing on upregulated/responsive gene set from WT cells treated with 2µM compared with ctrl **(c)** and on upregulated/responsive gene set from *Ezh2*$^{Y641F/WT}$ cells treated with 2µM compared with ctrl **(d).** Given the very limited number of significantly downregulated genes in mutant cells upon treatment, we didn't perform GO term analysis on this subset of genes.

Table S4a

	EZH2 mutant (n=30)	*EZH2* WT (n=130)
Age at diagnosis -yr		
Median	61,3	63,3
Range	37-85	28-87
Gender -no. (%)		
Female	21 (70)	80 (61,5)
Male	9 (30)	50 (38,5)
Death -no. (%)*	12 (40)	40 (30,7)
Histological Transformation -no. (%)*	4 (13)	30 (23)

* last update for the mutant cases, whether included in the restricted cohort or not, is June 2021, for the WT cases non included in the restricted cohort is December 2019

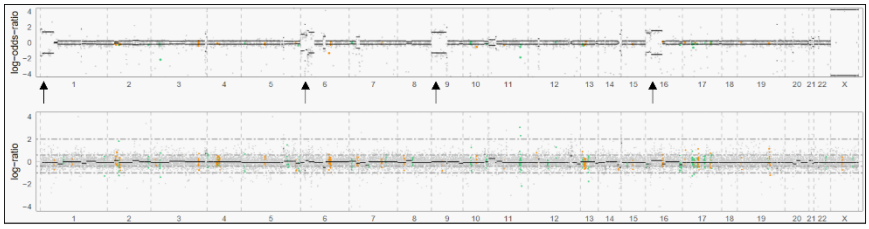
Table S4b

Number of *EZH2* mutations detected using Sanger sequencing (n = 160)	
EXON16	
Y646N	14 (46,7%)
Y646F	7 (23,3%)
Y646S	5 (16,7%)
Y646H	3 (10%)
EXON 18	
A692V	1 (3,3%)
TOTAL	**30 (18,8%)**

Table S4c

	EZH2 mutant (n=21)	*EZH2* WT (n=11)
Age at diagnosis -yr		
Median	62,4	53,5
Range	37-84	35-75
Gender -no. (%)		
Female	16 (76)	7 (64)
Male	5 (24)	4 (36)
Follow-up -yr		
Median	10	19
Range	1-21	8-24
Death -no. (%)	12 (57)	3 (27)
Histological Transformation -no. (%)	4 (19)	3 (27)
Ann Arbor stage at diagnosis -no. (%)		
I	4 (19)	1 (9)
II	3 (14)	1 (9)
III	4 (19)	1 (9)
IV	10 (48)	7 (64)
Missing data		1 (9)
FLIPI1 risk status -no. (%)		
Low	8 (38)	6 (55)
Intermediate	4 (19)	1 (9)
High	4 (19)	2 (18)
Missing data	6 (24)	2 (18)
FLIPI2 risk status -no. (%)		
Low	7 (33)	4 (36)
Intermediate	3 (14)	1 (9)
High	1 (5)	2 (18)
Missing data	10 (48)	4 (36)
Bone marrow involvment - no./total no. (%)	7/18 (39)	5/10 (50)
Extranodal involvment -no./total no. (%)	8/21 (38)	6/10 (60)
Bulk disease -no./total no. (%)	4/21 (19)	2/9 (22)

Figure S4a



Extended data S4

Extended data S4

Table S4a Elementary demographic and disease characteristics of patients from the initial cohort (n = 160) between *EZH2* mutant and WT patients. Patients with no death/histological transformation reported include lost of follow-up patients.

Table S4b Number and type of *EZH2* mutations detected throughout the initial cohort of 160 cases using Sanger sequencing of exons 16 and 18.

Table S4c Demographic and disease characteristics of patients from the restricted cohort (n = 32) between *EZH2* mutant and WT patients, including staging and prognostic factors. The risk groups according to the Follicular Lymphoma International Prognostic Index (FLIPI) 1 and 2 are based on the number of risk factors: zero or one risk factor indicates low risk, two risk factors intermediate risk, and more than two risk factors high risk. Patient with bone marrow involvement were classified as having extranodal involvement. Bulk disease was defined as a tumor that was 7 cm or larger in the greatest dimension.

Figure S4a Example of one allele-ratio plot (upper plot) and Copy Number Variation scatter plot (lower plot) by chromosomal position in one sample (P#70, inferred tumor cellularity: 0.58, inferred ploidy: 2.06). Arrows show regions with Copy number Neutral Loss of Heterozygosity.

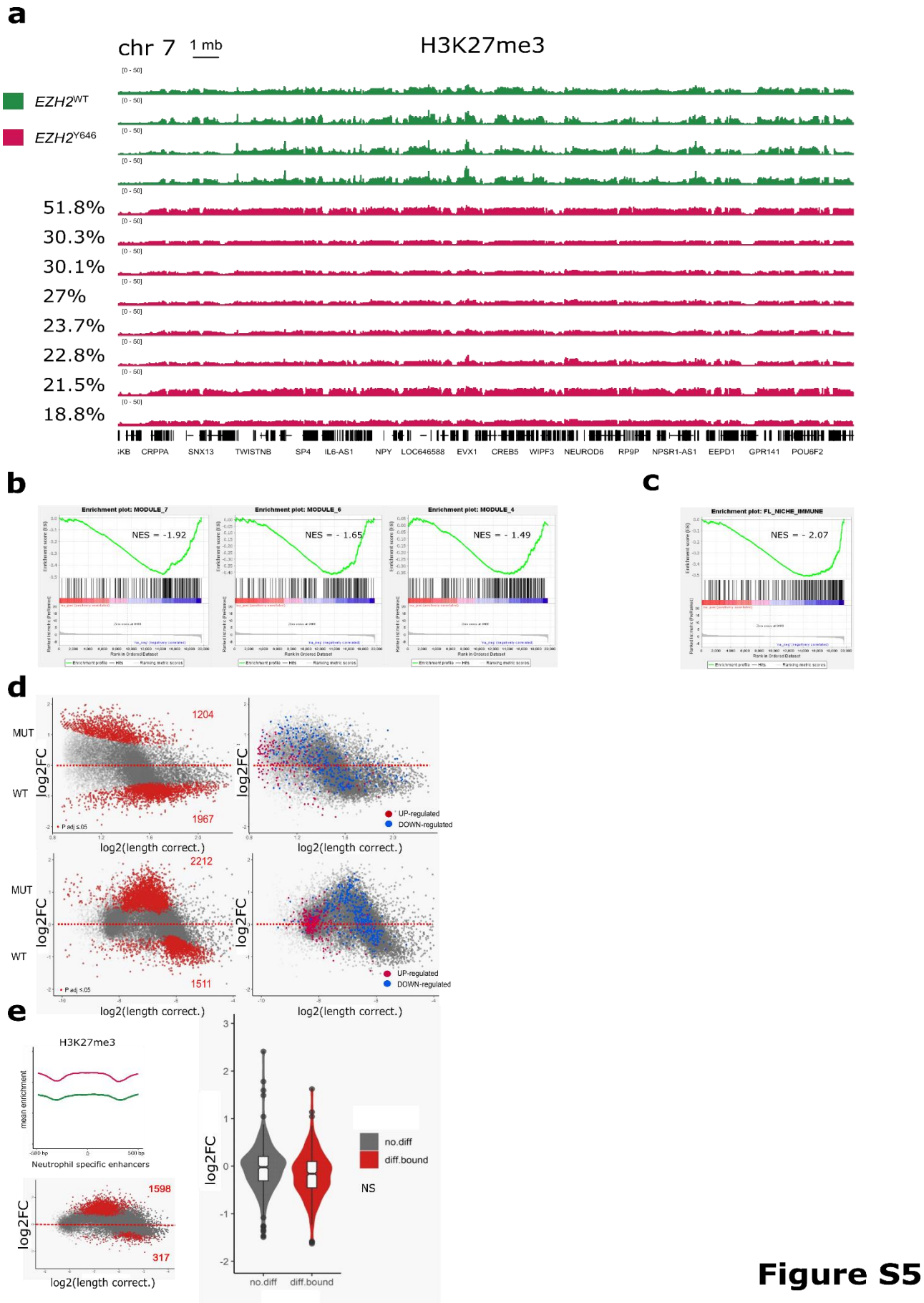

a
chr 7
1 mb
H3K27me3
EZH2WT
EZH2Y646
51.8%
30.3%
30.1%
27%
23.7%
22.8%
21.5%
18.8%
CRPPA
SNX13
TWISTNB
SP4
IL6-AS1
NPY
LOC646588
EVX1
CREB5
WIPF3
NEUROD6
RP9P
NPSR1-AS1
EEPD1
GPR141
POU6F2
b
Enrichment plot: MODULE_7
NES = -1.92
Enrichment plot: MODULE_6
NES = - 1.65
Enrichment plot: MODULE_4
NES = - 1.49
c
Enrichment plot: FL_NICHE_IMMUNE
NES = - 2.07
d
MUT
WT
log2FC
1204
1967
2212
1511
P adj ≤.05
log2(length correct.)
UP-regulated
DOWN-regulated
e
H3K27me3
mean enrichment
Neutrophil specific enhancers
1598
317
no.diff
diff.bound
NS

Figure S5

Figure S5

All the analyses were conducted in 4 *EZH2* WT and 8 *EZH2*^{Y6416} FL samples, after merge of 2 replicates per patient for ChIP-Seq experiments.

(a) Representative sampling of ChIP-Seq genomic tracks showing H3K27me3 distribution and in 4 WT and 8 *EZH2*^{Y6416} tumor sample. For the mutant cases, Variant Allele Frequency (VAF) of *EZH2* mutation assessed by DRAGON is reported on the left side of the track.

(b) Gene Set Enrichment Analysis of downregulated genes in *EZH2*^{Y6416} vs. *EZH2* WT FL samples in modules 7, 6 and 4 as defined by Béguelin *et al* (Cancer cell 2018, Mutant EZH2 induces a pre-malignant lymphoma niche by reprogramming the immune response). NES = Normalized Enrichment Score, nominal p value<1%.

(c) Similar as (b), in a set of 258 genes determining immune niche in FL as defined by Pangault *et al* (Cancers (Basel) 2020, Intergrative analysis of cell crosstalk within follicular lymphoma cell niche: towards a definition of the FL supportive synapse).

(d) MA-plot showing the enrichment in H3K27me3 according to the genotype (y axis, WT: log2FC<0, *EZH2*^{Y646}: log2FC>0) along the promoters (upper panels) and the gene bodies (lower panels). Base line enrichment in H3K27me3 in WT samples normalized on the length of the enhancer serves as a reference (x axis, log2 read counts/length of the enhancer). Significantly differentially bound gene bodies or promoters are represented by red dots (padj<0.05). On the right panels, downregulated genes are represented by blue dots and upregulated genes are represented by red dots.

(e) Similar legend as ***Fig5e*****,** with analyses performed on neutrophil specific enhancers.

Table S5a (next page)

Up (green) and down (red) regulated genes upon DESeq2 analysis in 8 *EZH2*^{Y641} and 4 *EZH2* WT FL

Table S5a

	log2FC	padj	pvalue
TRPT1	0.62319055859032	5.57136105142703e-07	1.37806321882447e-10
TRAK1	0.923183018759138	5.57136105142703e-07	3.34352229590675e-11
BCORP1	20.8648120147523	5.57136105142703e-07	1.01726261404408e-10
MIR378I	1.15870656594806	5.57136105142703e-07	1.41283183329792e-10
DNASE1	1.09952156039718	8.21033915551245e-07	2.49845488325175e-10
SPATS2	0.867748589828772	2.27944554938772e-06	8.09256927814272e-10
KDM1B	0.905542784022471	7.05619760829554e-06	3.22086415147638e-09
POM121B	2.07264461634053	1.38400231181202e-05	7.72126866659846e-09
KRT8P46	1.29850535422266	1.93450754229477e-05	1.17736422922033e-08
RP11-245C23.3	1.79447463844803	3.33265963300703e-05	2.19732085150334e-08
ATF7	0.941588119938736	9.5198733223256e-05	7.72521038480548e-08
TMEM131	0.82423427975599	0.000108071904438727	1.00227991125071e-07
CEP170	0.78100646587938	0.000108071904438727	1.08889774935213e-07
LRCH3	0.610571198575196	0.000108071904438727	1.15104224436439e-07
CTA-250D10.23	0.965169516096253	0.000108071904438727	1.13380066422281e-07
TNFRSF13C	1.16790651081073	0.000128158388470249	1.55997429795911e-07
TEX9	2.00944968270143	0.000144975281511389	1.83820157112376e-07
EIF4G3	0.854037685141104	0.000170430267739879	2.50670881191687e-07
BLK	1.05159088698647	0.000170430267739879	2.47017067470868e-07
RABGAP1L	0.83425332665067	0.000202034154498435	3.27894352282291e-07
CPT1B	1.08103998702283	0.000239297214937515	4.12644180548538e-07
STAT6	0.696965162023129	0.000325649061219174	6.11097797084214e-07
SDCCAG8	0.616352015393061	0.000347705177120885	6.98811435722981e-07
EP300	0.749989163387353	0.000347705177120885	6.81775144059514e-07
SNX29	0.785419232133387	0.000360628576578442	7.86480133533145e-07
SGK494	1.30780755572193	0.000395772179639011	8.83196018872876e-07
SLX4	0.929759402911331	0.00043334104616889	1.01098991346396e-06
SPRED2	1.18496264020778	0.000464466140502983	1.10716176921642e-06
TMEM79	0.699052600901259	0.000478110556857051	1.16393501694672e-06
NLK	1.62208722068135	0.000545502131799332	1.3972969690351e-06
PRKDC	0.509046109531442	0.000566410200152386	1.52253084181551e-06
SEPT7P2	0.705922540053133	0.000598052795993026	1.63791910450999e-06
MTOR	0.468410482903242	0.000606099900649396	1.69069810497118e-06
AF131215.5	1.88007522426312	0.000804192652446996	2.36563239042074e-06
RP11-480A16.1	1.10369262528225	0.000804192652446996	2.3436417180437e-06
SLC7A6	0.600583949355215	0.000816559487017295	2.48483893193882e-06
HYPK	1.02611760267867	0.000821179184299895	2.54054522707783e-06
XKR6	1.18610812519344	0.000852972657102255	2.72182736839183e-06
ATP6V0A2	0.594344531217668	0.000852972657102255	2.76868945856592e-06
GTPBP1	0.586007700223448	0.000930248256455703	3.07697672394596e-06
MICAL3	1.17690921488145	0.000930248256455703	3.11388065760899e-06
SMG7	0.473372194555096	0.000936263358125786	3.18150048153511e-06

EBF1	1.34013403936453	0.00101292487088987	3.54474900296196e-06
ASL	0.789763162484347	0.00114171092934569	4.13966926971901e-06
KIAA0319L	0.434613269210861	0.00114171092934569	4.22705775940737e-06
RP11-206L10.2	1.38301365989589	0.00114171092934569	4.19555636412679e-06
IGF2BP3	1.3323297729825	0.00115334768620642	4.32863664752626e-06
LRRFIP1P1	1.63255465179603	0.00116739850369085	4.49978629002913e-06
EDC4	0.65800486012247	0.00131957101827547	5.26984391150301e-06
WDFY4	0.936644009481292	0.00131957101827547	5.24265651080851e-06
ZNF274	0.661313500647405	0.00131957101827547	5.28711824535997e-06
AF131215.3	1.87778213220975	0.00144801319491514	5.87518667105602e-06
CBL	0.736675630576829	0.00156757206421942	6.43978988698956e-06
DTX2P1	1.34793545809678	0.00186197107635534	8.40469776312954e-06
RP11-175P13.3	1.15968381917547	0.00198485090286773	9.462696397503e-06
PDPK1	0.458761172994033	0.00203382015723404	1.00056071030938e-05
KMT2D	0.966013790789694	0.00203382015723404	9.90662301214415e-06
AKAP13	0.660665025437666	0.00203873265533979	1.02222907477203e-05
ZNF609	0.935465008881624	0.00203873265533979	1.0291845207096e-05
RP11-242D8.2	1.33629509368743	0.00203873265533979	1.04433736465648e-05
NCOR1	0.614159369906661	0.00222620982860888	1.1629538588361e-05
PAX5	0.665322719636942	0.00222620982860888	1.15928449511175e-05
PFKM	0.802444225438103	0.00223545827695729	1.18688012901094e-05
TTI1	0.37650072068339	0.0022633383708531	1.21730625586624e-05
PCDHGC4	2.19461834987999	0.00242776410545859	1.34212247043154e-05
KIAA0100	0.388004011275912	0.0024858728459895	1.41206957828688e-05
ANKRD17	0.428497851076673	0.00264954366704187	1.54535437292598e-05
RNU6-638P	1.60355080539273	0.00271740473431211	1.612498625118e-05
MANBA	0.558089952962898	0.00288994038655363	1.7588519875561e-05
SON	0.420310908678589	0.00327861493433299	2.02866065825747e-05
LRCH1	0.871348635866802	0.00329157202724902	2.08208088605997e-05
CPSF7	0.526461768247978	0.00329157202724902	2.08676017348546e-05
CUL9	0.593848019953881	0.00330322663389246	2.11090204326444e-05
RP5-1050D4.5	1.04099529814849	0.00334052563942789	2.15168005379795e-05
LRRC37A15P	1.18360548302518	0.00334828649983081	2.17366065820533e-05
KMT2C	0.860533041630093	0.00355367693421556	2.36106749699365e-05
FAM120A	0.38805020356292	0.00372494078268846	2.62599162546886e-05
RPAP1	0.651616579204934	0.00372494078268846	2.62491437632939e-05
ATAT1	0.682304816593105	0.00372494078268846	2.52548363580905e-05
E2F4	0.397284885396461	0.00372494078268846	2.58558261135747e-05
RP11-1407O15.2	0.620829033758855	0.00376115014867956	2.7468966952876e-05
ACSF3	0.902343479735027	0.00376115014867956	2.67972360522344e-05
SIPA1L3	1.09247596948219	0.00384309611352773	2.82623592058386e-05
NDUFA13	0.916280474419271	0.00400404900052549	3.00552443108877e-05
DENND3	1.06137826485156	0.00404125712898934	3.05394995293103e-05
CDK10	0.632686389610469	0.00405995132866966	3.08866815083658e-05
DOCK2	0.600326488105668	0.00430729052449138	3.32052624497991e-05
AC127496.1	1.12625079112441	0.00458648315914317	3.60554288008922e-05

RP11-92K15.3	1.78056478384702	0.00470766958108162	3.81984148726232e-05
NFYC	0.543719478504608	0.00472667979068204	3.85959043617086e-05
GLYR1	0.510237759157824	0.00477550412791264	3.92367839286832e-05
CRTC2	0.512017018023582	0.00479742223942215	3.96601828384546e-05
RP11-22B23.1	1.04527226576897	0.00490532914260083	4.08010335946917e-05
CLEC16A	0.733324775074453	0.00507720789763467	4.36279180553122e-05
RP11-252A24.2	0.617132675536533	0.00507720789763467	4.41072575348229e-05
MIR3916	0.665109744819289	0.00507720789763467	4.37681131315262e-05
GS1-124K5.2	1.3768617137334	0.0053361691594399	4.70910094711438e-05
IRF5	0.614295218799257	0.00536357505001161	4.83531845723235e-05
PAN2	0.872238113055372	0.00536357505001161	4.80602261641849e-05
AL133458.1	0.985269358378908	0.00536357505001161	4.80810450321083e-05
PCDHGA6	1.40251960596814	0.0053983356337912	4.92823661856477e-05
ABCC10	0.548586114399163	0.00558838176683761	5.15841903719858e-05
RNU4-78P	2.35430639397524	0.00567723209134775	5.32681410406925e-05
BCL6	1.03972347563487	0.00599220600886547	5.83363323480969e-05
RAF1	0.311534577756659	0.00599220600886547	5.83508421008353e-05
ILDR1	1.67113563232758	0.00599220600886547	5.78031560686526e-05
EIF4ENIF1	0.481124811320116	0.00599220600886547	5.70194071613943e-05
CHP1	0.494079545145221	0.00599220600886547	5.80251199770929e-05
SLC23A3	1.24722836283535	0.00599220600886547	5.74969421294989e-05
AF131215.2	1.31054564366308	0.00599220600886547	5.82762775104131e-05
DEPDC5	0.862752510545814	0.00624121081276139	6.14086776728564e-05
KLHL18	0.652554155647366	0.00649232422820536	6.42087145356822e-05
HIVEP1	0.841524939896981	0.00651799216809455	6.47931462670047e-05
SLC37A1	0.590792788050731	0.0067011915690229	6.76865079524534e-05
AC003102.3	1.52801087257106	0.00674249778885119	6.88669798329123e-05
RRNAD1	0.513823466238279	0.00674806183132523	6.95638097274562e-05
RNPS1	0.339783695437238	0.00674806183132523	7.05026493509661e-05
CTD-2540L5.6	1.93321921193557	0.00674806183132523	7.0299732260825e-05
RP3-418A9.3	1.88442817795018	0.00692277961588188	7.33016198980039e-05
LINC00843	0.793021694653693	0.00693405937253354	7.42043174724642e-05
ACACA	0.459111396406763	0.00716427519880596	7.74000228041747e-05
KCNAB3	0.818893959893465	0.00726433327686235	7.92124387343615e-05
CTC-325J23.2	1.29040509213317	0.00731371165072946	8.01218094313315e-05
PISD	0.531828021237429	0.00732477790413162	8.06145359434276e-05
GTF2I	0.662334523427334	0.0073347082777074	8.10958261672776e-05
AF131215.4	2.55052712426203	0.0080910591006956	9.10998184487712e-05
AL353898.2	1.1127538041207	0.00816692337681592	9.23682057630445e-05
RP11-384K6.6	0.929347427826199	0.00871090554081019	0.000100729647679907
RAET1K	1.37013539952311	0.00879675282201786	0.000102975412665641
C2CD3	0.540038768651279	0.00895426068281861	0.000106722689073509
GOLGA2B	1.00614547976374	0.00907693101173873	0.000109565835613624
KCNMB3	1.46663113944667	0.00937815003870253	0.000114628704129802
DIDO1	0.396657891487718	0.00941253283102422	0.000119544951615832
FANCD2	0.786338057084448	0.00941253283102422	0.000117843088501947

TLR10	0.800592283352203	0.00941253283102422	0.000116052345686833
RP11-295P9.3	0.663541100299504	0.00941253283102422	0.000118171854130362
RP11-506M12.1	1.05762156045441	0.00941253283102422	0.000116459453632618
RP11-1167A19.2	2.1243656527012	0.00941253283102422	0.000117704316649631
AP4B1	0.647678415758949	0.00945575140610027	0.000122567315557541
SIPA1L1	0.916490067143047	0.00945575140610027	0.000124447474266862
TAP2	8.15521889018799	0.00945575140610027	0.000124689119317648
RP11-552C15.1	2.27213785851536	0.00945575140610027	0.000122587326017277
WNK1	0.557400386334857	0.0096448119334459	0.000127671345266997
ARHGEF7	0.702075576587183	0.00975313390129722	0.000130083041151159
HNRNPL	0.340196921467996	0.00989991673567653	0.000134080897163647
TMC6	0.804372606338729	0.00989991673567653	0.000134562950000574
PHF12	0.515249796620436	0.0100678921825887	0.000137356747837721
BARX2	2.36349659034831	0.010102176055995	0.000139361560441783
ANKRD13A	0.686118541336334	0.010102176055995	0.000138633350975224
KB-1507C5.4	1.18421884052918	0.010102176055995	0.000138882157972112
SMYD4	0.482972197699649	0.0102436267497501	0.000142871499527377
MAST3	0.794393473730601	0.010379595625847	0.000145820763217508
TRIM41	0.681592658178966	0.0108992114951724	0.000155879058481362
AL135998.1	2.91772775912171	0.0108992114951724	0.000155719546342712
HAND2-AS1	2.1405624485045	0.0108992114951724	0.000154034230075472
CLCN6	0.804298718557886	0.01092273086045	0.000157258565331362
HECTD4	0.857486821006044	0.01092273086045	0.000159451856229625
RP11-313P13.3	1.58371666423567	0.01092273086045	0.000160652835093092
GS1-124K5.12	0.693142542811123	0.01092273086045	0.000159681628773209
ZDHHC20-IT1	1.30300757875961	0.010968544446033	0.000162216719164927
CACNG5	2.16562343540771	0.0110005506489832	0.000166260794917938
TTC13	0.438916498246139	0.0110005506489832	0.000165904230532619
SETD5	0.41203873147565	0.0110005506489832	0.000164165287890733
SPAG1	1.01049366516186	0.011407476991139	0.000173537729569625
PCDHGA4	1.87678003537237	0.011407476991139	0.000173568144106188
POLR3E	0.34091606565869	0.0114404028165277	0.000174649350701163
RP11-121A14.3	2.66494787864204	0.0115374535788193	0.000177886386770861
TMEM138	0.550245927418504	0.0116078899582956	0.000179561111593049
HTT	0.614247736550834	0.0117278884709286	0.000183201787748949
TMPPE	1.0083489169215	0.0119853282006026	0.000189654734421465
GIT2	0.483575956718441	0.0120159552163006	0.000192921361717652
NEK8	0.731895573465619	0.0120159552163006	0.000192175085318057
RP11-504P24.2	1.18311522825083	0.0120159552163006	0.000193781587082865
MROH1	1.04555743220387	0.0120471363328566	0.000194909798152926
VPS39	0.45782698064041	0.0121194858109027	0.000196695007328136
HAUS5	0.530921298707073	0.0123609474738202	0.000201565220564236
KPNA6	0.378540757593882	0.0124760176876548	0.000205221822865778
AMPD3	0.703768593702907	0.0124760176876548	0.000205769650439575
AC093642.3	1.45487490180446	0.0124760176876548	0.000206277920889357
NEK9	0.347800432877199	0.0125583165399687	0.000209549431538759

SREBF2	0.980233148140675	0.0125583165399687	0.000209055395557825
GS1-124K5.3	1.27040527583083	0.0125708434268081	0.000210396020228569
RP11-764K9.4	1.08465743510838	0.0126244010019707	0.00021385352409891
EFCAB4B	0.628687849093139	0.0129298955549045	0.00022296315304902
ERICH1	0.577604579868646	0.0129423993097418	0.000224341902544073
AC004893.11	0.677805886424424	0.0129796855585987	0.000225796629639364
VPS52	0.505866721242176	0.0130183100434981	0.000227789063498852
RBM33	0.470424664418084	0.0131618910355855	0.000230968925207313
INPP5D	0.714419538140437	0.0134448048399828	0.000240624647805263
ANKFY1	0.437292678619627	0.0134448048399828	0.000242070584682958
TARS2	0.48668728274831	0.0137118986460535	0.000251052158605534
DDB1	0.316807635008649	0.0137118986460535	0.000250653287529038
ZNF608	0.946135454683281	0.0137118986460535	0.000250991183420221
ATF6B	0.468633273953198	0.0137118986460535	0.000248275469459
RP11-242D8.1	0.793935199924434	0.0137118986460535	0.000249269884338798
CCNK	0.450679932076598	0.0138813944781648	0.000256971597328708
UBE3B	0.461778191431237	0.0138813944781648	0.000256059203029516
LA16c-431H6.6	1.68689257327701	0.0139061098818653	0.000258134412778957
AMH	2.50709612271842	0.0139766005865384	0.000260151768284202
SOX9	1.94331932372047	0.0140386017853994	0.000262729830035623
AC009121.1	1.80262398170877	0.0140958172434077	0.000265230420312637
TTC17	0.356100444516578	0.0143772823203815	0.000271984901633224
L3MBTL1	0.736455693028251	0.0143772823203815	0.000271908764036354
FGD2	0.543735022372757	0.0143975004531577	0.000276980652396683
SNX30	1.25316281251673	0.0143975004531577	0.000273483326962178
FBXW8	1.19583747244694	0.0143975004531577	0.000276095030116255
RP4-800G7.2	1.22451780877585	0.0144110770964918	0.000278471388840259
TNRC6C-AS1	0.653791248813638	0.0145615959341601	0.000282856988526821
ZNF384	0.452793935236715	0.0148270638702757	0.00028876566040401
TMEM134	0.584216277227364	0.0149181887150107	0.000292386856914668
RP11-458F8.1	1.69901307625543	0.0149181887150107	0.000292810216194611
MIR4520A	1.58231568145597	0.0149181887150107	0.000291829728473082
R3HDM2	0.584077962106372	0.0151741839423602	0.000300194515510902
RP11-458F8.2	1.37935355931485	0.0151741839423602	0.000299522933409966
NEURL4	0.576745257009317	0.0152768619251467	0.000303724191036035
ARSG	0.547985937947933	0.0155574089324703	0.000310879906648744
LRRK1	0.826120528808085	0.0157705109664235	0.000319149140841314
ZNF439	0.70244338741682	0.0157965567866291	0.000322869218695113
ZZEF1	0.492685566856683	0.0160001901870593	0.000327842817648321
MGEA5	0.49443328604891	0.0160144152158557	0.000328972877640314
LINC00664	1.63493226384277	0.0164263399190914	0.000339383457759688
SPEN	0.889173637725022	0.0167689570433087	0.000351753062157062
AC090587.2	0.924319015370914	0.0170809117970583	0.000360382375999202
CDK12	0.345419661624846	0.0172088809653148	0.000364827927347039
YEATS2	0.521853447858696	0.0175086888360387	0.00037511307503806
RANBP10	0.661395317868356	0.0177882112331774	0.000383424951772602

CYTH1	0.539419517380442	0.0177911723229049	0.000384391104608078
TRANK1	0.682417604059651	0.0179374777243332	0.000388461885088516
HERC2	0.700776209015673	0.0181756268671036	0.000397306648056075
RP11-644N4.1	1.60543379635813	0.0181756268671036	0.000397051344667599
INPP5B	0.850007559663866	0.0184581038084988	0.000405322992812854
AC138035.2	1.27803614312356	0.0184581038084988	0.00040535370234214
MAP4K1	0.468696701569004	0.0184851437412357	0.000406956104208682
PPME1	0.40755654232742	0.018522883564576	0.000409594625660857
VPS13B	0.485599570860178	0.0185573478167809	0.000411297915298132
DNAJC4	0.497468214416177	0.0189017890264751	0.000420849286535607
INO80E	0.550012081424095	0.0193178622517582	0.000433052447901664
SDHAP3	0.874772718355294	0.0195876816962446	0.000440993508842243
RP11-84E24.3	1.91943030485219	0.020033131929962	0.000454166961134707
RP11-175K6.1	0.997609758731243	0.0200705832213868	0.000456946060760661
ZNF611	0.453112658841633	0.0205997158991842	0.000475370022525171
BFSP2-AS1	1.30917717674715	0.020755940575689	0.000483424829153923
TUBGCP6	0.659989454527892	0.0211863796656524	0.000497504376182839
DGKD	1.05369229087115	0.021383743985422	0.000503223472599068
PLCG2	0.805799341708639	0.021450703219097	0.0005058871530598
RPL32P3	0.540261230140345	0.0223167260134852	0.000529706739073443
TRAF5	0.620566414485602	0.0225967688630434	0.000542084073247429
TFEB	0.472641612368137	0.022899830902789	0.000555161493712691
ACLY	0.425815437810339	0.022899830902789	0.000554706992974649
CTD-2547E10.2	0.870373944525245	0.0232632129574271	0.000568690401454576
PRRC2C	0.499582170843068	0.0233604398751072	0.000572251988622852
CTB-181H17.1	2.64683926212963	0.0233845216305882	0.000574027918507109
RP5-1107A17.4	2.02640415590336	0.0235128233483353	0.000580754930802824
FCHSD1	0.534823060699496	0.0235951507302776	0.000583985066509889
RP11-572O17.1	0.681689051108587	0.0239631366333911	0.000596738859207538
PRKAB1	0.621143440566154	0.0245523056174813	0.00061460823703283
MAN2B2	0.626300796394188	0.0245959368372029	0.000617486876016404
RNF123	0.51249914025878	0.0247603961070917	0.000624953911606313
DENND4B	0.636093548954277	0.0247603961070917	0.000625383032983296
ABCC1	0.663768435679316	0.0248586438542293	0.000629125286973699
FAM167A	1.8612985020579	0.025243068172634	0.000645255685905946
C17orf85	0.455389178136955	0.0252487671330923	0.000665890293107877
ZMYND8	0.483801807592811	0.0252487671330923	0.000660459586996948
UNC119	0.5583683929342	0.0252487671330923	0.000646790145329576
SORL1	0.862097910991093	0.0252487671330923	0.00065068733450858
POGZ	0.522272880324734	0.0252487671330923	0.000649856878870178
AMZ1	1.09314412253832	0.0252487671330923	0.000658253475598955
H3F3C	2.91419670210562	0.0252487671330923	0.000656778761802934
CTC-281F24.1	1.12393511156495	0.0252487671330923	0.000650247385306411
AL591479.1	1.31342877923801	0.0252487671330923	0.000664097204625509
IPO9	0.357788561517722	0.0252746006723515	0.000667853474174323
LAT2	0.783520376359853	0.0253425216237763	0.000670933523741504

TPCN1	1.0880318498935	0.0256388552970835	0.000681379529120646
AFG3L1P	0.556303499296132	0.0260724682867776	0.000698192587889565
RP5-1071N3.1	0.634910598177693	0.0261559756283611	0.00070415554276026
RP11-148O21.3	1.00707789255567	0.0261559756283611	0.000705735103427909
VMAC	0.547942448737503	0.0263245006870197	0.000712498715928592
CTC-429P9.5	0.73354304826955	0.0267652186712669	0.000727554601423724
LINC00174	1.20685778853505	0.0268741721077733	0.000736805576327115
RP4-555D20.2	1.87421094341411	0.0268741721077733	0.00073419546301105
SNRPN	0.712842362429246	0.0269745914751177	0.000741503706421554
EPS15L1	0.46661284898963	0.0269747805736382	0.000752465125706705
MYO19	0.494994672561363	0.0269747805736382	0.000753002234922834
DPAGT1	0.485223760522737	0.0269747805736382	0.000751857658533041
PATL1	0.419008967406111	0.0270600105006106	0.000757575990076433
RRN3P2	0.809325756531775	0.0277730595762244	0.00078600934404841
NUDT17	0.796158958120563	0.0277730595762244	0.000787398706857505
POLD4	0.638058492919494	0.0277773161248992	0.000789147454429618
AC018766.4	0.648387835212489	0.0277773161248992	0.000790336985650374
GGA3	0.513742171305568	0.028078705587977	0.000800726464382022
U62631.5	1.38731122153805	0.0285179523760697	0.000818641834196655
TP73	2.21539931175332	0.0285693130740641	0.000825431374679387
ATXN7	0.455471442129994	0.0286196604998416	0.000833167303844064
C19orf68	0.799488342591768	0.0286196604998416	0.000834625185748792
TNRC6B	0.49741339413408	0.0286426748838528	0.000838797753405507
RP11-148O21.2	1.26321717449898	0.0286426748838528	0.000839654414102903
PCDHGB3	1.17246318812584	0.0286426748838528	0.000839286809965039
PIK3C2B	0.714053366917336	0.0291362524145056	0.00085560126530399
KIAA0247	0.461691819860487	0.0292400163292919	0.00086013133189579
USP10	0.317481793946304	0.0293585300117251	0.000866595550379065
SUGP2	0.377277651089502	0.0294539672276264	0.000872400307396349
CHFR	0.435058504613631	0.0296984285915422	0.000882653504825467
CDK5RAP2	0.411898492873717	0.0296985551769365	0.000885361761156904
TADA2B	0.401893522317944	0.0297943148085105	0.000895358664692759
TLR6	0.712046290015091	0.0297943148085105	0.000896080979938466
NOM1	0.489706588130912	0.0298131505242665	0.000898714401990547
RP11-16E23.4	2.18804327973689	0.0298131505242665	0.00089967158096762
FKBP15	0.526726788884888	0.0298215407597848	0.000901437251753905
AL049542.1	2.06394854157022	0.0302052238244332	0.000917532986565275
RPS24P17	1.96723965037758	0.0302554676813674	0.000920691819689629
YJEFN3	0.864140774808927	0.0303372064871472	0.000927795075911639
ZKSCAN1	0.447033944476213	0.0303742060598963	0.000933323566984874
RNF31	0.360665836409699	0.0303784507897065	0.000935219335058671
RP11-138I18.2	1.56120229959347	0.0305070822164303	0.000946915571154606
HPS4	0.437600526790925	0.0306397123885996	0.000954537609262924
MED25	0.814127914613366	0.0306397123885996	0.00095378329264621
LRRC4	3.13702442115783	0.0306594931087117	0.000957866194399067
CMTR1	0.352250051885326	0.0307124760137451	0.000961079155068251

NBR2	0.732911946695944	0.0307705825470156	0.000964458082571164
PACS1	0.924653528506982	0.0309001623267086	0.000973133098777366
RMND5B	0.406207354637037	0.0311663080944325	0.000984764920770473
SSH2	0.577426139534606	0.0312273311391333	0.000988548044768351
RP11-234B24.4	2.47519393756124	0.0312599625887221	0.000992480427070044
AC009120.3	0.739340747250446	0.0314546994974155	0.00100025848683266
CYFIP2	0.607668339753804	0.0317276795435129	0.00101406394304503
CTD-2006C1.12	0.865958008663088	0.032323799363658	0.00104756848371342
RFX7	0.726904956302087	0.0323280724634362	0.00104934657283558
TBX10	4.68610468930066	0.0326954008896892	0.00107619390259209
ZNF808	0.554305420633353	0.0326954008896892	0.00107459525965949
AC074289.1	1.25195160156203	0.0326954008896892	0.00106937092506244
PCDHGC5	1.67964264140708	0.0326954008896892	0.00107150991370067
CTD-2196E14.6	0.730627357317897	0.0326954008896892	0.00107137286364316
ZNF137P	0.598109526753209	0.0327118002554036	0.00107839276593865
GOLGA3	0.631528101611326	0.0327164746560044	0.00108186547018892
ARHGAP33	0.960563811541587	0.0327610209196911	0.00108832320852045
MAP2K1	0.649744071434387	0.0327610209196911	0.0010876434745653
SYNE1	0.522888989181346	0.0331148908357814	0.00110343780895209
EPB41L2	1.25412841380488	0.0331875425678034	0.00111174487584704
PCSK9	2.35308445629718	0.0331875425678034	0.0011119418694227
CLK2	0.546433936728135	0.0331875425678034	0.00111174227151971
FLII	0.430851375696771	0.0332447579999376	0.00111788175452445
RHOA-IT1	1.0350668618259	0.0335718120408433	0.00113058189354973
AC007461.1	2.27563863832901	0.0336977061943332	0.0011382397081415
TNPO3	0.488224546840249	0.0342987621592852	0.00116411024620352
RP11-148O21.4	1.14741061279872	0.0349062280004401	0.00119322400326097
TANGO2	0.538393030437155	0.0350479643691723	0.00119984662723494
SLC25A35	0.912358741877344	0.0352729899069	0.00121219199822055
RP11-563N6.6	1.32663184183281	0.0352729899069	0.00121125124886513
CXXC1	0.384206617065691	0.0357542940076886	0.00123490765426971
TCOF1	0.456775599440279	0.0361039477386723	0.00125247757028209
PCYOX1L	0.715515156310278	0.0361039477386723	0.00125246130183855
SMCR8	0.672768580848879	0.0361208243448458	0.00126405481553703
ACAP2-IT1	1.25135936711151	0.0361208243448458	0.00125892141265574
CHD8	0.381302763022596	0.0364403763932196	0.00127893393843424
RP11-400F19.18	1.57718653138365	0.036472203938908	0.00128560033156875
ATXN2	0.508995740930254	0.0365192185460587	0.00129087578387745
MARS	0.43390026672226	0.0367888756103151	0.00130982353697019
SMAP2	0.679531606635658	0.0369926887562423	0.00131922030242139
AC131097.3	1.19379901959413	0.0369926887562423	0.00132083242300525
PHLPP1	1.48245747028932	0.0371679549959833	0.00133274555876453
STAG3L3	0.643836997900379	0.0374209418007211	0.00134751071047887
TXLNA	0.271380069538802	0.037456433638425	0.00135091800252566
PRPF8	0.518660749579544	0.037456433638425	0.00135448786246371
RP5-827C21.2	2.4212593650239	0.0377833362238117	0.00137014177613356

KDM4B	0.635310691312055	0.0381830114112713	0.00140461103110645
TEP1	0.441882084138758	0.0381830114112713	0.00140203978254783
SMARCC2	0.600348190949056	0.0381830114112713	0.00139436633901564
WDTC1	0.693772782226152	0.0381830114112713	0.00140829756931535
RP11-375N15.2	1.43293075393458	0.0387810600853274	0.00143779251013716
INPPL1	0.563109764568403	0.0388162220281647	0.00145264662487943
PPIEL	1.40184018837769	0.0388162220281647	0.00144556797511959
LETM1	0.349805720276237	0.0392540918036638	0.00147523872934599
RP13-638C3.2	1.61315185953062	0.0393826979715951	0.00148206937642256
LA16c-306A4.2	1.53204395125278	0.0395275942464675	0.00148952693234901
ADNP	0.336976109601306	0.0397623610565762	0.00151945171446899
ASH1L	0.568950706869106	0.0397623610565762	0.00151799794482217
MKL1	0.684848063890011	0.0397623610565762	0.00152240291438884
HMGB3P4	1.37459983329579	0.0397623610565762	0.00150597160763303
RP11-730K11.1	1.44221331101269	0.0397623610565762	0.00151979868924784
CTD-2574D22.4	0.74486363673267	0.0397623610565762	0.00151224964243531
RP11-274B21.1	1.47935649604969	0.0399458498174924	0.00153770350517202
CD22	1.07191506745859	0.0400257860000768	0.00154998241881975
MLYCD	0.450552486065285	0.0400257860000768	0.00155296070852862
AC010525.7	1.66721487540136	0.0400257860000768	0.00155226811987983
RP11-73M18.7	0.60402124584322	0.0400257860000768	0.0015433977419182
FAM21C	0.383543136344198	0.0400661838874925	0.00155859223217055
PAQR4	0.743132943262237	0.040345890732417	0.00157765794769455
SHISA8	1.77730690090224	0.040361074404348	0.00158029869859292
ATPAF2	0.411630820898789	0.0403776149049047	0.001584461110403
PUM1	0.336504964366625	0.040477362688322	0.0015930635211309
UBAP2L	0.41710432515897	0.0405095073401563	0.00159638318219311
TRIM60P18	0.688183497656001	0.040575987148614	0.00160106091198568
APBB3	0.680569306968704	0.040744614237199	0.0016135443588689
DOK3	1.29366621378817	0.040744614237199	0.00161598054133436
CTB-102L5.8	1.06774809043779	0.040744614237199	0.00161516670275623
MRPS25	0.551704091481321	0.0411595906862365	0.00163870156153044
ZNF585B	0.553911224856394	0.0416622972983061	0.0016783015844632
RLTPR	0.674583301959125	0.0417630328685373	0.00169215798948056
WASH5P	1.65876524770289	0.0417630328685373	0.00169718282878562
RP11-428G5.5	1.54214366515983	0.0417630328685373	0.00169430596150146
TPD52	0.909666891007615	0.0418161339695058	0.00170643359512416
FIZ1	0.325246206982881	0.0418646878631693	0.00171136270313508
L3MBTL2	0.55262128587904	0.0420050055370743	0.00171922906468626
MNT	0.626869188906232	0.0423341216346528	0.00174128785543964
DEF8	1.19537081113017	0.0423341216346528	0.00173956172345879
TNKS	0.994330353033879	0.0427683200724688	0.00176782374900147
MTMR4	0.44859198708082	0.0430824871370017	0.00178517989506164
RP11-108K14.4	0.805034962833474	0.0431816183163655	0.00179798869821932
RP11-408O19.5	1.87108781102203	0.0431816183163655	0.00179804780837531
CMIP	0.761456227134212	0.0432907892702488	0.00181137602819675

CYB5RL	0.738118500404859	0.0435490657235181	0.00182439155488289
KMO	1.08336885601373	0.0441597255323209	0.00185771139357632
TEC	1.41714273980064	0.0441597255323209	0.00185893250453042
STAG3L5P	1.04553986061336	0.0444243982672127	0.00187683338015865
RP4-739H11.4	0.778249993132429	0.0444251693345831	0.00187911909646712
PLAGL2	0.480079753603363	0.0445624116323729	0.0018894444451318
GPATCH8	0.448492363025706	0.0445624116323729	0.00188823708528009
ACD	0.466751529438586	0.0446065606914868	0.00189357870359459
ELMSAN1	0.695013224679617	0.0453888445833397	0.0019359952474813
NBPF15	0.526903061371137	0.0453888445833397	0.0019340745240916
UBE4B	0.298788773220439	0.0455735915787675	0.001948498133636
ZNF736	0.747333305458778	0.045723395281781	0.00195722197179201
RN7SL32P	1.31163826503875	0.0457366449372157	0.00196010878794681
RP11-261C10.3	1.37182130308138	0.0460252810565549	0.00198181587548893
CNDP2	0.42667737649801	0.0460877270901729	0.00199151714159493
AC005220.3	1.49418998870237	0.0462398819441778	0.00200448979486146
RP11-436H11.2	1.85431699430918	0.0463373424445097	0.00201170386633364
PRR14L	0.433507394958599	0.0468889101174548	0.00204040598878005
RP11-156E8.1	1.01051868752518	0.046931855934374	0.00205417617646522
AC016629.8	0.863177483284681	0.0477706079958376	0.00210300186338627
DDX23	0.367990368808585	0.0478063383569962	0.00210699944374041
KLC2	0.577058344599981	0.0480433371716932	0.0021271914262255
ZNF418	1.45394585529622	0.0485668797330927	0.00216268805627912
ITPKB-IT1	1.1857171278839	0.0486135694623075	0.0021672327208687
RN7SL431P	1.52167835405103	0.0486172317580959	0.00217479324494804
SLC25A44	0.465286629692134	0.048717259374177	0.00218694167502151
QSOX2	0.561516457090941	0.048717259374177	0.00219119702853443
AC005152.3	2.23644428037501	0.048717259374177	0.00218686876133885
CUBNP2	1.59501661532026	0.048717259374177	0.00218735242325095
GRAMD4	0.906442250707048	0.048732414728364	0.0022036233028685
WDR11	0.458429627610852	0.048732414728364	0.00220466165936505
CTC-425O23.2	0.844858228977943	0.048732414728364	0.00219575220740292
RP4-756H11.3	1.49993173746926	0.0489496921334282	0.00222875098485281
RP1-179N16.6	1.04799094700437	0.049341709730575	0.00225975370932237
PNKP	0.601760875132773	0.0493903632056171	0.002268828871547
SYVN1	0.712808634983331	0.0494246538307911	0.00227357919686197
RP11-75C10.7	1.67823914673788	0.0494369198593473	0.00227744110424079
BRPF1	0.305214564217172	0.0495401936663437	0.00229899463431072
AFMID	0.895734938261298	0.0498755564774912	0.00232467598533319
BRIP1	1.09522940263124	0.0499154646854318	0.00233413087386358
RP11-163E9.2	1.14700912184479	0.0499154646854318	0.00233233829880764
RP11-90B9.2	0.971887433437704	0.0499154646854318	0.00233109352036208

	log2FC	padj	pvalue
ANKS6	-1.12010298789152	5.57136105142703e-07	9.85103119965078e-11
RP11-67L2.2	-0.618811855986287	2.41117234896251e-06	9.78312055165597e-10
SORD	-1.07867712323701	1.2100657726073e-05	6.13716981593192e-09
ST7	-1.27431675525251	7.98048419177453e-05	5.66652019500144e-08
STK3	-0.713706845943246	8.15035586586907e-05	6.20050403144677e-08
SLC44A1	-1.012067933059	0.000106212704680936	9.15766079817368e-08
CASP1	-0.906020125297924	0.000110510797482295	1.23306666562382e-07
CD300E	-1.34765192100252	0.000117768841863796	1.37378067802774e-07
PRCP	-0.88327617823326	0.000159425749852478	2.10228203893311e-07
CD101	-2.11350531009869	0.000170430267739879	2.37990018439779e-07
FMO5	-1.55232736518178	0.000202034154498435	3.10592926938757e-07
SNCG	-1.51975259121309	0.000202034154498435	3.21904540157888e-07
MORC4	-1.03583858886117	0.000239297214937515	4.11069771748644e-07
CCR2	-1.35037348571318	0.000289006557884694	5.1302071947884e-07
TEX2	-0.655810769250875	0.000301600259576343	5.50672482870029e-07
RIMS1	-5.05086763338114	0.000347705177120885	7.05391646033139e-07
APBB2	-0.773769850699558	0.000354623805575745	7.3741319818459e-07
EMILIN3	-3.35768879733994	0.000357890390912591	7.62357174941869e-07
BIVM	-0.869726385072607	0.00043334104616889	9.90937660975439e-07
PLA2G4A	-0.917997514086894	0.000538822690298891	1.3390633374573e-06
SRD5A3	-1.23831873552885	0.000545502131799332	1.43866261873334e-06
RBMS1	-0.53773045530332	0.000545502131799332	1.41333232278278e-06
GLIPR2	-0.746530547018704	0.000752592190754448	2.13750381306736e-06
CLEC2B	-0.997040792826406	0.000816559487017295	2.46964040840702e-06
MEGF9	-0.877705750619578	0.000852972657102255	2.73874266856953e-06
GGTA1P	-1.08465654127807	0.00101292487088987	3.51140369570983e-06
RHOC	-0.640778306282763	0.00102609362602832	3.64287436334037e-06
BX255923.3	-2.6705519829264	0.00115674087276657	4.40003882220889e-06
UAP1	-0.752356747202497	0.00160530500184115	6.67621900649055e-06
MNDA	-1.82266259191762	0.00175295679668236	7.44258841205801e-06
RAP2B	-0.671136035273394	0.00175295679668236	7.46809204855294e-06
BMPR2	-0.82739629657085	0.00181250661844846	7.81371722717043e-06
SCARB2	-0.486545449998014	0.00186197107635534	8.32990939003023e-06
FAM105A	-1.10210364194986	0.00186197107635534	8.31643713970072e-06
HCAR3	-2.3293108903738	0.00186197107635534	8.15133097334267e-06
DSTN	-0.622978348764328	0.0019168871758865	8.7827213115867e-06
LRRC28	-0.801941016675961	0.0019168871758865	8.84702201174983e-06
USP53	-2.99318312251222	0.00193102924800529	9.01022928521004e-06
EIF2AK4	-0.75089920660382	0.00198485090286773	9.40942246829968e-06
PLSCR1	-0.532762667357318	0.00203382015723404	9.96736109377001e-06
SLC11A1	-1.48614953035499	0.00203873265533979	1.03774294192793e-05
IFIT2	-0.907911085198031	0.00223545827695729	1.19046061307762e-05
PTPN4	-0.715030429983854	0.0022633383708531	1.22826599219598e-05
GNAQ	-0.646198126618108	0.00238723924435699	1.30761190034263e-05
ACER3	-0.603539033390056	0.0024575307426111	1.3835061745186e-05

TBC1D9	-0.992883449504583	0.0024575307426111	1.38195943390304e-05
S100A9	-1.67462825136204	0.00257375694258878	1.47504455298743e-05
WDSUB1	-1.02524024446633	0.00264954366704187	1.54393252961545e-05
PAK1	-0.916114051851443	0.00271405513143775	1.59674593115981e-05
ICAM2	-0.701964542537627	0.0027174806089087	1.6263260731918e-05
RP11-420G6.4	-1.18163971535888	0.00285814459238144	1.72500485111016e-05
EHBP1	-0.682745817772656	0.00289326017976494	1.77554639017882e-05
CCDC112	-0.997964492141117	0.00329157202724902	2.0686348112555e-05
CCL20	-3.08782782845928	0.00345907026020297	2.26312351557632e-05
SESN1	-0.964901413947143	0.00355367693421556	2.35789127934743e-05
ARHGAP10	-1.09835822681026	0.00372494078268846	2.53364294038552e-05
CYSTM1	-0.519916257257048	0.00372494078268846	2.55044689350539e-05
TPPP3	-1.45607349405686	0.00372494078268846	2.58896299684685e-05
RP11-64C12.3	-3.34914676091963	0.00372494078268846	2.59414005648439e-05
KLF3	-0.626960540713677	0.00376115014867956	2.74439735696924e-05
CLIC2	-0.527788162307243	0.00376115014867956	2.74465109585332e-05
TCEAL1	-1.23050028867714	0.00376115014867956	2.69263948373769e-05
AC005550.4	-1.45105136194387	0.00395215168553403	2.92648042850316e-05
VAMP5	-0.799291454545131	0.00398744166518303	2.97283524259221e-05
PROS1	-0.968726091854125	0.00414046395408899	3.17091878616137e-05
IGLV10-54	-6.08425728799365	0.00431497679535228	3.34833620575594e-05
LEPROTL1	-0.748172554379375	0.00434619856525735	3.39460657833155e-05
PTGER4	-0.766898689849512	0.00459678155079255	3.63695248731368e-05
CHN2	-2.14776090201323	0.00470766958108162	3.75552102306242e-05
RNLS	-0.919296100263588	0.00470766958108162	3.81601262827701e-05
AKR1C1	-1.11619580863737	0.00470766958108162	3.82019137279028e-05
NENF	-0.832018031876266	0.00500526447101496	4.18861204908185e-05
ST3GAL4	-0.671275931420297	0.00507720789763467	4.30085026563553e-05
DIAPH2	-0.621793683176551	0.00507720789763467	4.42476181007634e-05
HDAC11	-0.587595724611799	0.00507720789763467	4.32587828090964e-05
TCEAL3	-0.887010381561274	0.00507720789763467	4.42907013436711e-05
FCGR1A	-1.72623347935841	0.00510078659472271	4.47550885472957e-05
CTNNAL1	-1.04554786644594	0.00536357505001161	4.84209747376409e-05
DNAJC6	-2.09667099653978	0.00538330387847799	4.88721100698666e-05
DOCK7	-0.890132538148764	0.0055247079242504	5.0716241532146e-05
OAT	-0.766642429000542	0.00563752782027428	5.2323760770411e-05
TM4SF1	-1.04428023220147	0.00567723209134775	5.31479500489562e-05
ITM2B	-0.466719683010615	0.00624121081276139	6.1283609724885e-05
MEFV	-1.69749485657118	0.00663810112471067	6.6323777530456e-05
GUCY1A3	-0.76950676625096	0.0067011915690229	6.79737441702379e-05
CERS6	-1.00584009873247	0.0067011915690229	6.7651925981205e-05
MT1E	-0.907240874996247	0.00674249778885119	6.90766624409363e-05
WWP1	-0.619007764415219	0.00674806183132523	7.0332764236938e-05
MYRFL	-2.74580246525773	0.00681318864109253	7.15286325864053e-05
MFAP3L	-1.07217603236224	0.00692277961588188	7.34179881912778e-05
RP11-47L3.1	-0.996014930079432	0.00692277961588188	7.37325008538415e-05

OSM	-1.60542370215096	0.00716427519880596	7.73016683681372e-05
IFI44	-1.54858265678068	0.00716427519880596	7.77580206189824e-05
SPG20	-0.633669999484044	0.00757023116115646	8.40838172284457e-05
MET	-1.22363457961317	0.00797652618359302	8.9001154353627e-05
DDX60	-0.933356597630595	0.0080910591006956	9.08261620768228e-05
TCEAL4	-0.983146599422925	0.0082548465487167	9.3781286550314e-05
APOL6	-0.69312567371264	0.00829821888017864	9.46948951686461e-05
EPHA4	-1.94795490239688	0.00835692308827552	9.57886401557168e-05
CPNE8	-0.5457129846935	0.00870022565898672	0.000100164894486483
BLVRA	-0.588598475804845	0.00875737553612371	0.000101711162842843
HERC5	-1.53861500424741	0.00879675282201786	0.000103060805492018
UXS1	-0.664628774087499	0.00886330453816778	0.000104739562681599
AC020571.3	-1.5531130414423	0.00886330453816778	0.000104395314616254
GALNT12	-0.793609561319058	0.00895426068281861	0.000106434258386719
RP11-588K22.2	-0.78067375845947	0.00900359625045879	0.000108014119106504
IGHV3OR16-13	-4.31046518353958	0.00900359625045879	0.000108223984955051
FPR2	-1.62623886329448	0.00915042325625755	0.000110917033942565
ACVR1B	-0.500262050655197	0.00931181697535153	0.000113345644574954
GNPTAB	-0.478990930540598	0.00941253283102422	0.000118058805770381
DUSP5	-1.62973727815804	0.00941253283102422	0.000120300160948324
RFESD	-0.924467788082776	0.00941253283102422	0.000119828465671596
PLA2G16	-0.855402965600035	0.00941253283102422	0.000118481025316658
RP11-385J1.2	-3.38434882130362	0.00941253283102422	0.000119492992964462
MPP5	-0.557575627970187	0.00945575140610027	0.000121774032830297
FLRT3	-2.9302678933067	0.00945575140610027	0.000124203803596638
GIMAP4	-0.873064430792653	0.00945575140610027	0.000124180108286065
RP11-122A3.2	-1.50881603307115	0.00945575140610027	0.000124413224796139
AOAH	-0.481633893854215	0.00975313390129722	0.000130094548665678
DNAJC12	-1.72999761226842	0.00989240612732173	0.000132813801076117
SLC19A2	-0.903062169747646	0.00989240612732173	0.000132955704404334
ULK2	-0.616502534697977	0.00989991673567653	0.000133849779186296
CD33	-0.842998443146784	0.0102239244334284	0.0001415596373853
RTN3	-0.450699426609069	0.0102370561171214	0.000142260657102565
CCL13	-1.74367638639673	0.0102626052096163	0.000143656694114424
POLR3G	-0.95360741391348	0.0108992114951724	0.0001558846498777
ADIRF	-1.38885555259599	0.0108992114951724	0.00015577848276839
PDE1A	-1.04510257567371	0.01092273086045	0.000160391733531223
TNFSF10	-0.541301001859886	0.01092273086045	0.000158467752223236
CARD6	-0.789759463160186	0.01092273086045	0.000159991004858637
RP11-422J8.1	-0.729578048240426	0.01092273086045	0.000157912149216658
KAZALD1	-1.06668701758378	0.010968544446033	0.000162439264504825
TRPC1	-0.872998463219842	0.0110005506489832	0.000165567359922583
AK4	-1.22284519766563	0.0110005506489832	0.000165612162952691
RP11-536O18.1	-1.08019252939018	0.0110005506489832	0.000165786980135336
TDO2	-1.30492818111661	0.0114733881633474	0.000175734808811224
PCBP4	-0.767945734779787	0.0115140835917296	0.000176942097088506

COQ3	-0.805161346202757	0.0116895879667012	0.000182010625641693
MCU	-0.572292986908805	0.0116895879667012	0.000181647301798159
USP46	-0.658280773996333	0.0117631532425573	0.00018434925962115
CAPN2	-0.800562603919052	0.0119853221706534	0.000188923962052831
MEI1	-0.815549967674754	0.0119853221706534	0.000189046771571395
SLK	-0.796655204133492	0.0120159552163006	0.00019312050489571
PRDX4	-1.0309455905426	0.0120159552163006	0.000193589512646124
RIMS2	-2.51821016214075	0.0120159552163006	0.000193795899923091
VPS54	-0.662144290809849	0.0123609474738202	0.000201867682029219
CTNNBIP1	-0.557740727443334	0.0124352096062184	0.000203711147882972
C5orf30	-0.798040693980475	0.0125583165399687	0.000208467845132902
RGS18	-0.886208192492658	0.0125950797498806	0.000211440452260003
CSF2	-2.49911288186892	0.0126173736661655	0.00021245463595714
S100A10	-0.796446583029237	0.012619040719853	0.000213122714394231
WNT5A	-1.35352047216902	0.0126401144930638	0.000214760782835948
CCT8P1	-1.48122263085227	0.0127371636638179	0.000217055687530699
IGSF6	-0.662273078379949	0.0127439236719152	0.000217817227642919
CD244	-0.935905972728351	0.0127991375355628	0.000219410076939708
EIF4E3	-0.751020163268817	0.0128541136244478	0.000221004438742598
EXOC6B	-0.83189389109461	0.0129423993097418	0.000224491584111767
BEX4	-1.13787271324312	0.0130183100434981	0.000227150618777477
HCAR2	-2.31755264988892	0.0131902359652185	0.000232135308613421
LGR4	-1.25806636320959	0.0132820957823508	0.000234425588692908
C6	-2.08251595659838	0.0133313261432891	0.000236104048264914
ARHGEF3	-0.775273699380141	0.0133313261432891	0.000236646759149526
TRBV30	-1.13434739499375	0.0134054369399197	0.000238642205503465
ARHGAP22	-1.34081352110146	0.0134448048399828	0.000241761591272423
EMR1	-1.28114291901257	0.0134448048399828	0.000241676558771622
MXI1	-0.724631120849001	0.0136699109277398	0.000246816873270547
SLC38A5	-1.47015096465047	0.0138813944781648	0.000256292820016796
RP11-247L20.4	-0.918919228973119	0.0138813944781648	0.000256127819045559
MLTK	-0.939585755110013	0.0140386017853994	0.0002620455390289
HADH	-0.678528644780834	0.0140958172434077	0.000264871318272525
OCRL	-0.769974860455859	0.0143975004531577	0.000276973872626319
AIG1	-0.559718313280571	0.0143975004531577	0.000277050457319046
CD1E	-2.12539319971378	0.0143975004531577	0.000274408550937507
IFIT1	-1.50331245350275	0.0143975004531577	0.00027747883411269
HPCAL4	-1.29324282621203	0.0144699792189007	0.000280343463083637
FCN1	-1.28340349346036	0.0150326754395868	0.000295819753033406
NKIRAS1	-0.856980103254058	0.0151741839423602	0.000300913218109389
LEPROT	-0.428625047991688	0.0154374774600727	0.000307700392646375
ENTHD1	-1.88174088138293	0.0156169519647536	0.000312861795713226
PKIA	-1.28201441526734	0.0157705109664235	0.000318080584347707
CYP4X1	-1.54374916461518	0.0157705109664235	0.000320737175916004
DPYD	-0.535026980269071	0.0157705109664235	0.000319739450534308
AL928768.3	-2.13167620318412	0.0157705109664235	0.000320445711759279

RP11-196O16.1	-2.09461337145593	0.0157705109664235	0.000319660122207016
RAB9B	-0.959160700062627	0.0157837508541264	0.000321806960661298
RYK	-0.546006430127099	0.0160144152158557	0.000329758714694802
CORO1C	-0.860564283570509	0.0164263399190914	0.00034061924448315
GIMAP5	-0.970392373836383	0.0164263399190914	0.000340740124101455
MANF	-0.523809009345597	0.0166769747834853	0.000347630807730003
CLGN	-1.49182472112632	0.0166769747834853	0.000347054057403718
REPS2	-0.691896770047225	0.0166781196861204	0.000348500548292418
MTUS2	-3.45054601339986	0.0167689570433087	0.00035209962042551
DDX60L	-1.01150138087025	0.0170126239330194	0.000358078761079427
SOCS6	-0.700295554056288	0.0171706505936123	0.000363146589112762
MYCT1	-1.00434395816042	0.0172760893125905	0.000367128945680145
GIMAP6	-0.730475374278131	0.0173260015722886	0.000369599583877235
SOX5	-2.82488746205381	0.0173260015722886	0.000369947084340087
GNAI1	-0.890823544415593	0.0175086888360387	0.00037562384630747
BCL2L14	-1.602504435218	0.0177292401547801	0.000381254644501028
VIMP	-0.414500194369108	0.0181402822561869	0.000393773941555409
GNLY	-1.49910258689123	0.018144592896032	0.000394787764487382
FCN3	-1.73665994434081	0.0184851437412357	0.000407822565676194
AC092580.4	-1.3910633520735	0.0186303150431216	0.000413860018709096
DYNC2LI1	-0.828999724199874	0.0190942551622334	0.00042610297060317
SFRP1	-1.30600041063523	0.0192428145458659	0.000430394137786014
HEPACAM2	-4.65763192209687	0.0195876816962446	0.000441087927835502
KLHL8	-0.628393796581586	0.019919835266012	0.000449577861407686
SOCS2	-0.692525620436577	0.0199511206072369	0.000451295825471809
MLLT3	-0.798250138290754	0.0200705832213868	0.000457051877385134
DNAJC3	-0.413181638557827	0.0203205258008238	0.000463774235957332
SLC18B1	-0.779793667670557	0.0203514508828875	0.000466544393115847
RP11-750B16.1	-1.83884765845292	0.0203514508828875	0.000466169906150293
MAGEH1	-0.855219697645801	0.0204678756755789	0.000470251441955534
FHL2	-1.79932311665501	0.020537399918101	0.00047289037697509
LAMP3	-1.84357280563252	0.0207273946257966	0.00047936764971158
PPAP2A	-0.656293681360929	0.020755940575689	0.000482849290429188
KL	-0.809211293235406	0.020755940575689	0.00048423860956621
ITGB1	-0.764686890321419	0.020755940575689	0.000482712642394059
TSHZ2	-1.16781906438298	0.0208058858233166	0.000486459063982805
KLRF1	-1.39229441861233	0.0211719400161613	0.000496091509228915
RHOU	-0.890224151938048	0.0215019564758888	0.000509142558693793
IDO1	-1.14034314360854	0.0215019564758888	0.000509276952591169
MGST2	-0.616095826268018	0.0224068207023856	0.000534425605751758
C12orf57	-0.631840273669189	0.0224068207023856	0.00053311462752345
MAN1C1	-0.846282594843577	0.0224068207023856	0.000535254478410693
IGKV2D-28	-3.23422852269062	0.0225499232144802	0.000539816592647697
LPCAT2	-0.622922937907485	0.022899830902789	0.00054117680968025
STAT1	-0.712962364222473	0.022899830902789	0.000551287162639198
TTC28-AS1	-0.637777358313698	0.022899830902789	0.000552975744992924

ZNF204P	-1.07989610291943	0.0229274172647695	0.000556993095796753
ODF2L	-0.670306517314559	0.0231497809847369	0.00056356924849996
TMEM55A	-0.71984190893643	0.023241605756182	0.000566983434027668
APOL3	-0.72534551737723	0.0234110177995446	0.000575865681025466
CLEC10A	-1.59445601139437	0.0234403180365502	0.000577775248048049
AAMDC	-0.653630475845162	0.0236537042249644	0.000586963698265368
FAM26F	-1.07724030739434	0.0236537042249644	0.000587833598936581
GPR126	-1.05952957012925	0.0245523056174813	0.000615113899376915
B2M	-0.567568853637967	0.0245523056174813	0.000615146268450361
NREP	-0.607560673681289	0.0247522278825013	0.000622665975032745
ENTPD6	-0.627806354046948	0.0250673587294001	0.000635678823588783
RNF149	-0.442672289777835	0.0251678935198112	0.000639504724523275
PLP2	-1.17575026819459	0.025210350247018	0.000643140750329667
IGHV3-72	-3.96788774956969	0.025210350247018	0.000642650092614386
PDE8A	-0.420641357785963	0.0252487671330923	0.00065254415380081
LIPA	-0.841654387317735	0.0252487671330923	0.00066410924076362
SWT1	-0.79311520158142	0.0252487671330923	0.000659983820797609
UCK2	-0.578302872951259	0.0252487671330923	0.00066466173447322
CLEC4D	-1.71032652996316	0.0252487671330923	0.000664419166464388
RP1-28O10.1	-1.51658658760622	0.0252487671330923	0.000655285216220584
CTD-2319I12.1	-1.14475791124415	0.0252487671330923	0.000665676451062087
SPARCL1	-0.79029199086083	0.0255303894276342	0.000677202093150718
CSRNP3	-1.39948423965017	0.0258459942465503	0.000688195312645884
STRADB	-1.20546182873726	0.0259475248650515	0.000692214742558052
LYPD1	-1.76724723663734	0.0260724682867776	0.000697201527390625
CREG1	-0.439524797529731	0.026134936237315	0.000701190914923144
PPT1	-0.393028878673382	0.0261559756283611	0.000705181457619185
SQRDL	-0.736129103856383	0.0263245006870197	0.000712952445446495
ARHGAP18	-0.560690557897049	0.0267652186712669	0.000727603449196076
PHF16	-1.22440713050211	0.0268741721077733	0.000737380286570238
IFI44L	-2.07929655614782	0.0268741721077733	0.000734327168704951
SCN11A	-1.25951480781148	0.0268741721077733	0.000736773762815793
FMO3	-1.06277239731466	0.0269747805736382	0.000746963620427572
ISM1	-1.76750111121718	0.0269747805736382	0.000749643138506946
GIMAP2	-0.690842867196495	0.0269747805736382	0.000753751197284535
CPD	-0.667532601183238	0.0269747805736382	0.000747115052004652
OPTN	-0.382323489879783	0.0269747805736382	0.00075382178303366
GBP4	-0.83593610846439	0.0269747805736382	0.00075317621155496
REPS1	-0.379334400043564	0.0270779846553981	0.000759452529007209
HRSP12	-0.886600535636229	0.0271031603159838	0.000761533236042756
IL1B	-1.78473445350894	0.0272163363999659	0.000766093558958315
AP3S1	-0.655110621318029	0.0275256203564549	0.000776195410974739
ACAT1	-0.803223603661755	0.0277571075224712	0.000784130896688971
PTPRO	-1.87864311791231	0.028078705587977	0.000801760472994424
RP11-383J24.6	-1.00946128008621	0.0281172503477289	0.000804287122590613
TSPAN2	-1.27295964474725	0.028403991396296	0.00081392986452844

RP11-169D4.1	-2.93777725337487	0.0285255024927268	0.000820305315888629
KITLG	-0.620021170312528	0.0285693130740641	0.000825912078521912
S1PR1	-1.0426087166066	0.0285693130740641	0.000823816381948395
GLCE	-0.715984540450477	0.0285872344708105	0.000827880046803916
PLS1	-1.50103104455357	0.0286196604998416	0.000832074740066146
RNF157-AS1	-1.88767513358984	0.0286196604998416	0.000833784303641926
TYROBP	-0.747961408101023	0.029261499738979	0.000862247367669868
CCDC104	-0.497297693322093	0.029409619618956	0.000869595183742523
TMEM66	-0.544961781789752	0.0296984285915422	0.000881759574534735
AC034220.3	-1.41979789940242	0.0296985551769365	0.00088566974915244
CCR7	-1.33797270402401	0.0297507625636271	0.000890436728433119
S100A8	-1.50432264141905	0.0297507625636271	0.000891264035333942
CD8A	-1.00762730957454	0.0297507625636271	0.000891753343566649
SLC4A5	-1.53594034466727	0.0300012694774363	0.000908391635544427
GPA33	-1.47545295316398	0.0302052238244332	0.000917630931218516
CARD17	-1.17773301007305	0.0302802824510524	0.00092298269275663
CD1B	-1.89645855443627	0.0303372064871472	0.000926787688338883
PGRMC1	-0.422217339513894	0.0303742060598963	0.000933548149936457
RP11-1094M14.12	-1.46511779620735	0.0303742060598963	0.000933276724740436
ACTG2	-2.65391950326406	0.0303883568081387	0.000937065524133911
DDO	-0.939604555287704	0.0304166277307745	0.000939479955776319
PLB1	-1.08295069768134	0.0304338865709765	0.000941556565821154
AXIN2	-1.07956586583323	0.0304797379274597	0.000944520965343505
TCF4	-0.938218518904736	0.0306397123885996	0.000955694229293946
RTP4	-0.753731522683215	0.0309001623267086	0.000970476701249207
SAP30	-0.937964777677625	0.0309001623267086	0.000973221119079273
TBX21	-1.07508758221026	0.0309496892953702	0.000976350699483708
CRYBG3	-0.595308467474314	0.0312273311391333	0.00098986062595518
C3orf52	-1.47607287843904	0.0317054655844945	0.00100984086763009
SYP	-0.868626625871131	0.0317276795435129	0.00102013563137833
DUSP3	-0.430217276441075	0.0317276795435129	0.00101564994407965
TPT1	-0.396945496320129	0.0317276795435129	0.00101894161390569
ARRB1	-0.590790520308413	0.0317276795435129	0.00101639119828624
GPR155	-0.798548594946053	0.0317276795435129	0.00102234813906024
ZNF449	-0.693186332494256	0.0317276795435129	0.00102342162548431
RP11-108M9.4	-1.41760908373355	0.0317276795435129	0.00102227277594883
STOM	-0.575879535641556	0.0319680614013671	0.00103279683078921
C6orf99	-2.98013206538085	0.032323799363658	0.00104613469230383
LILRA1	-1.17255777609303	0.0325474645667737	0.00105815600064941
TXN	-0.687576342460591	0.0325474645667737	0.00106020456847451
NDUFAF2	-0.685602748787196	0.0325474645667737	0.00106142007995311
GPR160	-0.830566335672168	0.0326954008896892	0.00107311575789339
IGLV4-69	-3.22002099589834	0.0327164746560044	0.00108102099482046
TRIM2	-0.612129507661539	0.0327610209196911	0.00108551285466155
RP11-1399P15.1	-1.56028890405437	0.0330762651285175	0.00110047319188048

SRC	-0.689552367318857	0.0331875425678034	0.00111259145089608
EEF1A1P24	-1.8004641446764	0.0332447579999376	0.00111720292951129
AGO4	-0.6610995399633	0.0336742267843545	0.00113573874380462
USP27X	-1.2920171110935	0.0337245341007991	0.00114085632932155
CD8B	-1.19290276661417	0.0339622162094741	0.00115061928426884
SLC16A5	-0.787144216298318	0.0342987621592852	0.00116585990696245
ABCD2	-0.739338600600078	0.0342987621592852	0.00116723991524473
MGP	-1.32554266390166	0.0344023773864407	0.00117251090955461
CD300LG	-1.6905289048674	0.0348253062513952	0.00118869154065979
CXCL11	-1.37104408331458	0.0352729899069	0.00121291713530853
SLFN11	-0.579799570437158	0.0354384080115641	0.00122040264948278
EIF5A2	-0.915938259023993	0.0355617114848007	0.00122645249326289
CPPED1	-0.684676084883729	0.0358003194061644	0.00123831302099732
PAIP2B	-1.39797433671902	0.0361208243448458	0.00126113205590519
TSPAN7	-1.02817325158559	0.0361208243448458	0.00125813331770243
RAVER2	-1.35536930197396	0.0361208243448458	0.00125927054971048
RGPD1	-2.04423472717865	0.0361208243448458	0.00126290777539765
C9orf89	-0.745936531065026	0.0363561992931593	0.00127413570581595
CMTM3	-0.604883174342505	0.036472203938908	0.001284552852314
ABHD6	-0.617465946510093	0.036472203938908	0.00128392504289283
PSTPIP2	-1.09006103727773	0.0365192185460587	0.00129096187688811
ACKR4	-1.61734930126832	0.0367662298004749	0.00130175213257339
AC007276.5	-1.46657244881705	0.0367662298004749	0.00130342316937323
TLR7	-1.02631968303924	0.0367703514647061	0.00130543419512574
SYDE2	-1.05390723672128	0.0367888756103151	0.00130834166742902
RP11-539L10.2	-1.27057065638882	0.0369975304694426	0.00132288172546315
EVA1C	-0.930668722626984	0.0370302532371862	0.00132592984660209
GMFG	-0.767382819938895	0.0372095810494315	0.00133612534274979
MCTP1	-0.633264267012966	0.0374209418007211	0.00134651433788031
LRRC16A	-0.891266788457971	0.037456433638425	0.00135440934258978
NMT2	-0.67656483781164	0.0375710908958798	0.00136053958004048
CRYGN	-2.14494872583542	0.0380339954065712	0.00138116045600776
NOS2	-1.57994634222067	0.0381830114112713	0.00140981043299719
ACVR2B	-0.789555345393821	0.0381830114112713	0.00139195230248946
MRAP2	-1.31087267682157	0.0381830114112713	0.00139441167077922
KCTD3	-0.504027247114739	0.0381830114112713	0.00139672407673976
MPC2	-0.602899901654724	0.0381830114112713	0.0014030956784696
FTH1	-0.307074141852808	0.0381830114112713	0.00140925958331324
GIMAP7	-0.919026491834619	0.0381830114112713	0.00140395478935878
C2orf82	-1.49162038858451	0.0381830114112713	0.00140332310031587
PERP	-0.882024516655778	0.0384562624926901	0.00142184994457428
IGKV2D-40	-2.45830158994785	0.0386101774990045	0.00142949888797856
FOCAD	-0.681573780644342	0.0388061442073405	0.00144069065069601
SGK3	-0.567901016924189	0.0388162220281647	0.00144668971017455
GBP1	-1.23379579879713	0.0388162220281647	0.00144940595232205
MACROD1	-0.806712860757225	0.0388162220281647	0.00145009075333834

IGKV6-21	-2.81313041719406	0.0388162220281647	0.00145287679955295
GPR171	-0.860521593746189	0.0388802354312809	0.00145724471185863
ATP7A	-0.649758810287456	0.0389284373045225	0.0014610256938351
ACVR2A	-0.656532135294243	0.0395676556003102	0.00149304335175893
ATG4A	-0.438579047421528	0.0397623610565762	0.00151545963258257
NOSIP	-0.569591296838853	0.0397623610565762	0.00150712619992849
GBP5	-1.06944174207949	0.0397623610565762	0.00152257354555536
NBL1	-1.06958801330759	0.0397623610565762	0.00151566398318918
ICA1L	-1.17651593603182	0.0397623610565762	0.0015068522394758
NEK3	-0.768536405032638	0.0397901626756642	0.00152565618414577
CSF3R	-0.84362128841721	0.0398563060785898	0.00153021370905779
BHLHB9	-0.843716713381516	0.0398790148988609	0.00153310814491741
LRRC8C	-0.715886473996687	0.0400257860000768	0.0015515244721815
AMER1	-0.661385419234173	0.0400257860000768	0.00154594097506893
RCN1	-0.532960738800513	0.0400661838874925	0.00155763414513612
TMEM71	-1.15037787655849	0.0401540346076408	0.00156608270088126
MAML2	-0.816646074239563	0.0401540346076408	0.00156488047811193
RDX	-0.557363186030781	0.040291106189835	0.00157347222022483
CLIC5	-1.04476578852102	0.0403776149049047	0.00158504204171001
NDFIP1	-0.437043168378124	0.0404756383402655	0.00159094282668285
PRSS12	-1.67596541487369	0.040744614237199	0.00161089287953882
KHK	-1.04431001492589	0.0410143216728328	0.00163083776393472
ABCA10	-1.087979978415	0.0410143216728328	0.00163026147142976
GOLGA7B	-1.08975178174663	0.0411923820075574	0.00164310580709372
IGHV3-49	-2.450109798913	0.0411923820075574	0.00164418545620265
PPP1R36	-1.17157871890948	0.0412317642040474	0.00164784856686054
MAPK8	-0.458495368532717	0.0412875488136856	0.00165217203499508
CHST10	-0.776958766966617	0.0413786449036987	0.00165791598488218
CTGF	-1.60003994447093	0.0413920699457889	0.00166055319405178
LHFP	-0.806114409249982	0.0416210594759552	0.00167185064182972
FKBP1B	-1.83637564284918	0.0416295363854035	0.00167430249802835
APMAP	-0.408305454696621	0.0416622972983061	0.00167984614049568
PTPN12	-0.745773253037901	0.0417630328685373	0.00169873471423477
NMNAT2	-1.13608413753817	0.0417630328685373	0.00169660618710026
TCEAL8	-0.531071044424202	0.0417630328685373	0.00168929400216466
IGHV3-43	-2.62794031137414	0.0417630328685373	0.0016931305530097
RAPGEF4	-0.817522176650079	0.0417956822598034	0.00170218252546646
LA16c-366D3.1	-1.60884699675402	0.0418161339695058	0.00170725707995396
MB21D2	-1.02201073308444	0.0420734575821716	0.00172416461563091
RP11-750H9.5	-0.618180742418367	0.0423234151680555	0.00173655438814003
PLCL2	-0.753826965263212	0.042458506862436	0.00174855746676969
PTGER2	-1.02038939965277	0.042473741527616	0.00175133904052096
SYTL2	-0.729668567111675	0.0427127458102948	0.0017633603027631
CREM	-0.748227950754117	0.0429212825352255	0.00177632330216281
CCL24	-1.42200440155354	0.0431816183163655	0.00179518162865493
C8orf31	-0.931901311538194	0.0431816183163655	0.0017974664980047

STX10	-0.297548118470067	0.0432762389657992	0.00180637747470978
NFAM1	-0.626676650558376	0.0432762389657992	0.0018044319215745
AQP9	-1.51440507917482	0.04328657842327	0.00180900444392019
FHOD3	-1.55714374740172	0.0438901981802319	0.00184090855074564
TSHZ1	-0.404822137613034	0.0441072715688552	0.00185225038591125
ATP11C	-0.526507760177556	0.0443077237586398	0.00186740977042297
IFI27	-0.882762050663748	0.0444243982672127	0.00187675296209091
RNF130	-0.520662289276141	0.0448867557991675	0.00190774972661675
ACO1	-0.502985340100162	0.0451692333049143	0.00192204629217544
S100A6	-0.636059737445136	0.0454692233895032	0.00194172978110066
SMAP1	-0.418744476055736	0.0457422281817796	0.00196266800435084
ICK	-0.468545790026748	0.0458061249606508	0.00196773281136437
UBE2E3	-0.58744102919964	0.0460252810565549	0.00197991250639914
CDC14A	-0.693898242193914	0.0460877270901729	0.0019882011830312
PPIB	-0.343843575146421	0.0460877270901729	0.00199056917751374
UTS2	-1.32365274098229	0.0461576118949265	0.00199687797060264
FAM171B	-0.845393328051554	0.0462398819441778	0.0020051275073425
DRP2	-1.74475944107358	0.0464865988382102	0.00202054142132912
TEAD2	-0.563743438154307	0.046897770277171	0.00204533915382832
ENPP6	-1.338499113848	0.046897770277171	0.00204554863510509
IL15	-0.675994397976842	0.046931855934374	0.00205044629958459
RP11-553P9.1	-2.03452459317043	0.046931855934374	0.00205249013168327
MTSS1	-1.00561350710563	0.0471057293417038	0.00206417559219111
HGF	-1.08290239247827	0.0477581705997511	0.00209518778560555
AGTPBP1	-0.576270700538136	0.0477706079958376	0.00210051511778217
SUSD4	-1.22823356355724	0.0477706079958376	0.00210115344413116
NADK2	-0.84258847846284	0.047851847347573	0.00211239312442302
GIMAP1	-0.871052168112209	0.047851847347573	0.00211385905765259
CYTH3	-0.739987622735155	0.0479965358635332	0.00212268495577425
NDUFB11	-0.651678585730162	0.0480994096690689	0.00213252347196326
OXNAD1	-0.609278236458389	0.0480994096690689	0.00213455309937796
CYYR1	-0.805988763750145	0.0484791361957267	0.00215632207960909
TPK1	-0.759378750330383	0.0484791361957267	0.00215465444992403
ARAF	-0.45272176350166	0.0486172317580959	0.00217240178174856
FAHD2CP	-0.869595976601556	0.0486172317580959	0.00217263059519732
CASP6	-0.691552066266357	0.048717259374177	0.00219162190317467
CD300LB	-1.65685973387676	0.048732414728364	0.00219855375235107
SLC22A5	-0.833969961537906	0.048732414728364	0.00220279342070651
FDX1	-0.740835950792427	0.0488332247106575	0.00221665243779675
FAM117B	-0.663158958948778	0.0488332247106575	0.00221602466786608
LINC00933	-1.591291974762	0.0488332247106575	0.00221172509204182
SUCNR1	-0.962726139604765	0.0489496921334282	0.00222889499155703
IGLV5-37	-3.36082969832923	0.0489496921334282	0.0022293870028817
RCAN1	-0.772975356012015	0.0490747937290224	0.00223757364519912
MPZL2	-0.822386193562783	0.0492016138626838	0.00224771866964616
MRGPRF	-1.17046072319187	0.0492016138626838	0.00224834681190232

PROCR	-0.835832187357198	0.0492850692118994	0.0022546600613244
NGEF	-3.08599964496962	0.0493903632056171	0.00226949683340717
LEPR	-1.058179552506	0.0493903632056171	0.00226827936416781
CARD16	-0.820826356560784	0.0494369198593473	0.00228166541928316
IGKV2-24	-3.34143794151729	0.0494369198593473	0.00228072934423503
TLR8	-1.26490924773346	0.0494622730607312	0.00229036137873143
CD63	-0.487519273072071	0.0494622730607312	0.00228857743848113
C20orf203	-1.31670771991735	0.0494622730607312	0.00228658950251069
RP9P	-1.07871371939215	0.0495401936663437	0.00229868780401283
AC093627.10	-0.95990357943112	0.0496128846434664	0.00230488422850409
FRZB	-1.02725046510725	0.0498755564774912	0.00232375456969576
TRIB1	-0.739068563778529	0.0498755564774912	0.00232332100324318

REFERENCES

1. Rd, K. Chromatin structure: a repeating unit of histones and DNA. Science (New York, N.Y.) 184, (1974).

2. Luger, K., Dechassa, M. L. & Tremethick, D. J. New insights into nucleosome and chromatin structure: an ordered state or a disordered affair? Nat Rev Mol Cell Biol 13, 436–447 (2012).

3. Fyodorov, D. V., Zhou, B.-R., Skoultchi, A. I. & Bai, Y. Emerging roles of linker histones in regulating chromatin structure and function. Nat Rev Mol Cell Biol 19, 192–206 (2018).

4. Trojer, P. & Reinberg, D. Facultative heterochromatin: is there a distinctive molecular signature? Mol Cell 28, 1–13 (2007).

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

6. Zuin, J. et al. Cohesin and CTCF differentially affect chromatin architecture and gene expression in human cells. Proc Natl Acad Sci U S A 111, 996–1001 (2014).

7. Cramer, P. Organization and regulation of gene transcription. Nature 573, 45–54 (2019).

8. Spitz, F. & Furlong, E. E. M. Transcription factors: from enhancer binding to developmental control. Nat Rev Genet 13, 613–626 (2012).

9. McKnight, S. L. & Kingsbury, R. Transcriptional control signals of a eukaryotic protein-coding gene. Science 217, 316–324 (1982).

10. Kassis, J. A., Kennison, J. A. & Tamkun, J. W. Polycomb and Trithorax Group Genes in Drosophila. Genetics 206, 1699–1725 (2017).

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

12. Laugesen, A. & Helin, K. Chromatin repressive complexes in stem cells, development, and cancer. Cell Stem Cell 14, 735–751 (2014).

13. Wang, H. et al. Role of histone H2A ubiquitination in Polycomb silencing. Nature 431, 873–878 (2004).

14. Cao, R. et al. Role of histone H3 lysine 27 methylation in Polycomb-group silencing. Science 298, 1039–1043 (2002).

15. Blackledge, N. P. et al. Variant PRC1 complex-dependent H2A ubiquitylation drives PRC2 recruitment and polycomb domain formation. Cell 157, 1445–1459 (2014).

16. Isono, K. et al. SAM domain polymerization links subnuclear clustering of PRC1 to gene silencing. Dev Cell 26, 565–577 (2013).

17. Illingworth, R. S. et al. The E3 ubiquitin ligase activity of RING1B is not essential for early mouse development. Genes Dev 29, 1897–1902 (2015).

18. Kaustov, L. et al. Recognition and specificity determinants of the human cbx chromodomains. J Biol Chem 286, 521–529 (2011).

19. Schuettengruber, B., Bourbon, H.-M., Di Croce, L. & Cavalli, G. Genome Regulation by Polycomb and Trithorax: 70 Years and Counting. Cell 171, 34–57 (2017).

20. Arrigoni, R. et al. The Polycomb-associated protein Rybp is a ubiquitin binding protein. FEBS Lett 580, 6233–6241 (2006).

21. Rose, N. R. et al. RYBP stimulates PRC1 to shape chromatin-based communication between Polycomb repressive complexes. Elife 5, (2016).

22. Farcas, A. M. et al. KDM2B links the Polycomb Repressive Complex 1 (PRC1) to recognition of CpG islands. Elife 1, e00205 (2012).

23. He, J. et al. Kdm2b maintains murine embryonic stem cell status by recruiting PRC1 complex to CpG islands of developmental genes. Nat Cell Biol 15, 373–384 (2013).

24. Wu, X., Johansen, J. V. & Helin, K. Fbxl10/Kdm2b recruits polycomb repressive complex 1 to CpG islands and regulates H2A ubiquitylation. Mol Cell 49, 1134–1146 (2013).

25. Fursova, N. A. et al. Synergy between Variant PRC1 Complexes Defines Polycomb-Mediated Gene Repression. Mol Cell 74, 1020-1036.e8 (2019).

26. Scelfo, A. et al. Functional Landscape of PCGF Proteins Reveals Both RING1A/B-Dependent- and RING1A/B-Independent-Specific Activities. Mol Cell 74, 1037-1052.e7 (2019).

27. Gao, Z. et al. PCGF homologs, CBX proteins, and RYBP define functionally distinct PRC1 family complexes. Mol Cell 45, 344–356 (2012).

28. Gao, Z. et al. An AUTS2-Polycomb complex activates gene expression in the CNS. Nature 516, 349–354 (2014).

29. Almeida, M. et al. PCGF3/5-PRC1 initiates Polycomb recruitment in X chromosome inactivation. Science 356, 1081–1084 (2017).

30. Zhao, W. et al. Essential Role for Polycomb Group Protein Pcgf6 in Embryonic Stem Cell Maintenance and a Noncanonical Polycomb Repressive Complex 1 (PRC1) Integrity. J Biol Chem 292, 2773–2784 (2017).

31. Trojer, P. et al. L3MBTL2 protein acts in concert with PcG protein mediated monoubiquitination of H2A to establish a repressive chromatin structure. Mol Cell 42, 438–450 (2011).

32. Endoh, M. et al. Polycomb group proteins Ring1A/B are functionally linked to the core transcriptional regulatory circuitry to maintain ES cell identity. Development 135, 1513–1524 (2008).

33. Kuzmichev, A., Nishioka, K., Erdjument-Bromage, H., Tempst, P. & Reinberg, D. Histone methyltransferase activity associated with a human multiprotein complex containing the Enhancer of Zeste protein. Genes Dev 16, 2893–2905 (2002).

34. Pasini, D., Bracken, A. P., Jensen, M. R., Lazzerini Denchi, E. & Helin, K. Suz12 is essential for mouse development and for EZH2 histone methyltransferase activity. EMBO J 23, 4061–4071 (2004).

35. Hauri, S. et al. A High-Density Map for Navigating the Human Polycomb Complexome. Cell Rep 17, 583–595 (2016).

36. Poepsel, S., Kasinath, V. & Nogales, E. Cryo-EM structures of PRC2 simultaneously engaged with two functionally distinct nucleosomes. Nat Struct Mol Biol 25, 154–162 (2018).

37. Margueron, R. et al. Ezh1 and Ezh2 maintain repressive chromatin through different mechanisms. Mol Cell 32, 503–518 (2008).

38. Shen, X. et al. EZH1 mediates methylation on histone H3 lysine 27 and complements EZH2 in maintaining stem cell identity and executing pluripotency. Mol Cell 32, 491–502 (2008).

39. Bracken, A. P. et al. EZH2 is downstream of the pRB-E2F pathway, essential for proliferation and amplified in cancer. EMBO J 22, 5323–5335 (2003).

40. Wassef, M. et al. EZH1/2 function mostly within canonical PRC2 and exhibit proliferation-dependent redundancy that shapes mutational signatures in cancer. Proc Natl Acad Sci U S A 116, 6075–6080 (2019).

41. Yi, Y. et al. A PRC2-independent function for EZH2 in regulating rRNA 2'-O methylation and IRES-dependent translation. Nat Cell Biol 23, 341–354 (2021).

42. Margueron, R. et al. Role of the polycomb protein Eed in the propagation of repressive histone marks. Nature 461, 762–767 (2009).

43. Montgomery, N. D. et al. The murine polycomb group protein Eed is required for global histone H3 lysine-27 methylation. Curr Biol 15, 942–947 (2005).

44. Rai, A. N. et al. Elements of the Polycomb Repressor SU(Z)12 Needed for Histone H3-K27 Methylation, the Interface with E(Z), and In Vivo Function. Mol Cell Biol 33, 4844–4856 (2013).

45. Grijzenhout, A. et al. Functional analysis of AEBP2, a PRC2 Polycomb protein, reveals a Trithorax phenotype in embryonic development and in ESCs. Development 143, 2716–2723 (2016).

46. Beringer, M. et al. EPOP Functionally Links Elongin and Polycomb in Pluripotent Stem Cells. Mol Cell 64, 645–658 (2016).

47. Zhang, Z. et al. PRC2 complexes with JARID2, MTF2, and esPRC2p48 in ES cells to modulate ES cell pluripotency and somatic cell reprogramming. Stem Cells 29, 229–240 (2011).

48. Chen, S., Jiao, L., Liu, X., Yang, X. & Liu, X. A Dimeric Structural Scaffold for PRC2-PCL Targeting to CpG Island Chromatin. Mol Cell 77, 1265-1278.e7 (2020).

49. Healy, E. et al. PRC2.1 and PRC2.2 Synergize to Coordinate H3K27 Trimethylation. Mol Cell 76, 437-452.e6 (2019).

50. Brien, G. L. et al. Polycomb PHF19 binds H3K36me3 and recruits PRC2 and demethylase NO66 to embryonic stem cell genes during differentiation. Nat Struct Mol Biol 19, 1273–1281 (2012).

51. Ballaré, C. et al. Phf19 links methylated Lys36 of histone H3 to regulation of Polycomb activity. Nat Struct Mol Biol 19, 1257–1265 (2012).

52. Cai, L. et al. An H3K36 methylation-engaging Tudor motif of polycomb-like proteins mediates PRC2 complex targeting. Mol Cell 49, 571–582 (2013).

53. Gatchalian, J., Kingsley, M. C., Moslet, S. D., Rosas Ospina, R. D. & Kutateladze, T. G. An aromatic cage is required but not sufficient for binding of Tudor domains of the Polycomblike protein family to H3K36me3. Epigenetics 10, 467–473 (2015).

54. Walker, E. et al. Polycomb-like 2 associates with PRC2 and regulates transcriptional networks during mouse embryonic stem cell self-renewal and differentiation. Cell Stem Cell 6, 153–166 (2010).

55. Li, H. et al. Polycomb-like proteins link the PRC2 complex to CpG islands. Nature 549, 287–291 (2017).

56. Liefke, R., Karwacki-Neisius, V. & Shi, Y. EPOP Interacts with Elongin BC and USP7 to Modulate the Chromatin Landscape. Mol Cell 64, 659–672 (2016).

57. Conway, E. et al. A Family of Vertebrate-Specific Polycombs Encoded by the LCOR/LCORL Genes Balance PRC2 Subtype Activities. Mol Cell 70, 408-421.e8 (2018).

58. Li, G. et al. Jarid2 and PRC2, partners in regulating gene expression. Genes Dev 24, 368–380 (2010).

59. Son, J., Shen, S. S., Margueron, R. & Reinberg, D. Nucleosome-binding activities within JARID2 and EZH1 regulate the function of PRC2 on chromatin. Genes Dev 27, 2663–2677 (2013).

60. Pasini, D. et al. JARID2 regulates binding of the Polycomb repressive complex 2 to target genes in ES cells. Nature 464, 306–310 (2010).

61. 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).

62. da Rocha, S. T. et al. Jarid2 Is Implicated in the Initial Xist-Induced Targeting of PRC2 to the Inactive X Chromosome. Mol Cell 53, 301–316 (2014).

63. Kasinath, V. et al. Structures of human PRC2 with its cofactors AEBP2 and JARID2. Science 359, 940–944 (2018).

64. Schmitges, F. W. et al. Histone methylation by PRC2 is inhibited by active chromatin marks. Mol Cell 42, 330–341 (2011).

65. Sengupta, A. K., Kuhrs, A. & Müller, J. General transcriptional silencing by a Polycomb response element in Drosophila. Development 131, 1959–1965 (2004).

66. Brown, J. L., Mucci, D., Whiteley, M., Dirksen, M. L. & Kassis, J. A. The Drosophila Polycomb group gene pleiohomeotic encodes a DNA binding protein with homology to the transcription factor YY1. Mol Cell 1, 1057–1064 (1998).

67. Klymenko, T. et al. A Polycomb group protein complex with sequence-specific DNA-binding and selective methyl-lysine-binding activities. Genes Dev 20, 1110–1122 (2006).

68. Scheuermann, J. C. et al. Histone H2A deubiquitinase activity of the Polycomb repressive complex PR-DUB. Nature 465, 243–247 (2010).

69. Ragazzini, R. et al. EZHIP constrains Polycomb Repressive Complex 2 activity in germ cells. Nat Commun 10, 3858 (2019).

70. Tanay, A., O'Donnell, A. H., Damelin, M. & Bestor, T. H. Hyperconserved CpG domains underlie Polycomb-binding sites. Proc Natl Acad Sci U S A 104, 5521–5526 (2007).

71. Mendenhall, E. M. et al. GC-rich sequence elements recruit PRC2 in mammalian ES cells. PLoS Genet 6, e1001244 (2010).

72. Bartke, T. et al. Nucleosome-interacting proteins regulated by DNA and histone methylation. Cell 143, 470–484 (2010).

73. Bracken, A. P., Dietrich, N., Pasini, D., Hansen, K. H. & Helin, K. Genome-wide mapping of Polycomb target genes unravels their roles in cell fate transitions. Genes Dev 20, 1123–1136 (2006).

74. Cooper, S. et al. Jarid2 binds mono-ubiquitylated H2A lysine 119 to mediate crosstalk between Polycomb complexes PRC1 and PRC2. Nat Commun 7, 13661 (2016).

75. Blackledge, N. P. et al. Variant PRC1 complex-dependent H2A ubiquitylation drives PRC2 recruitment and polycomb domain formation. Cell 157, 1445–1459 (2014).

76. Margueron, R. et al. Role of the polycomb protein EED in the propagation of repressive histone marks. Nature 461, 762–767 (2009).

77. Yuan, W. et al. H3K36 methylation antagonizes PRC2-mediated H3K27 methylation. J Biol Chem 286, 7983–7989 (2011).

78. Kaneko, S., Son, J., Bonasio, R., Shen, S. S. & Reinberg, D. Nascent RNA interaction keeps PRC2 activity poised and in check. Genes Dev 28, 1983–1988 (2014).

79. Zhang, Q. et al. RNA exploits an exposed regulatory site to inhibit the enzymatic activity of PRC2. Nat Struct Mol Biol 26, 237–247 (2019).

80. Beltran, M. et al. G-tract RNA removes Polycomb repressive complex 2 from genes. Nat Struct Mol Biol 26, 899–909 (2019).

81. Jermann, P., Hoerner, L., Burger, L. & Schübeler, D. Short sequences can efficiently recruit histone H3 lysine 27 trimethylation in the absence of enhancer activity and DNA methylation. Proc Natl Acad Sci U S A 111, E3415-3421 (2014).

82. Ferrari, K. J. et al. Polycomb-dependent H3K27me1 and H3K27me2 regulate active transcription and enhancer fidelity. Mol Cell 53, 49–62 (2014).

83. Conway, E., Healy, E. & Bracken, A. P. PRC2 mediated H3K27 methylations in cellular identity and cancer. Curr Opin Cell Biol 37, 42–48 (2015).

84. Jung, H. R., Pasini, D., Helin, K. & Jensen, O. N. Quantitative mass spectrometry of histones H3.2 and H3.3 in Suz12-deficient mouse embryonic stem cells reveals distinct, dynamic post-translational modifications at Lys-27 and Lys-36. Mol Cell Proteomics 9, 838–850 (2010).

85. Jung, H. R. et al. Precision mapping of coexisting modifications in histone H3 tails from embryonic stem cells by ETD-MS/MS. Anal Chem 85, 8232–8239 (2013).

86. Peters, A. H. F. M. et al. Partitioning and plasticity of repressive histone methylation states in mammalian chromatin. Mol Cell 12, 1577–1589 (2003).

87. 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 U S A 107, 20980–20985 (2010).

88. 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).

89. Hansen, K. H. et al. A model for transmission of the H3K27me3 epigenetic mark. Nat Cell Biol 10, 1291–1300 (2008).

90. Xu, C. et al. Binding of different histone marks differentially regulates the activity and specificity of polycomb repressive complex 2 (PRC2). Proc Natl Acad Sci U S A 107, 19266–19271 (2010).

91. Jiao, L. & Liu, X. Structural basis of histone H3K27 trimethylation by an active polycomb repressive complex 2. Science 350, aac4383 (2015).

92. Kanhere, A. et al. Short RNAs are transcribed from repressed polycomb target genes and interact with polycomb repressive complex-2. Mol Cell 38, 675–688 (2010).

93. Riising, E. M. et al. Gene silencing triggers polycomb repressive complex 2 recruitment to CpG islands genome wide. Mol Cell 55, 347–360 (2014).

94. Wassef, M. & Margueron, R. The Multiple Facets of PRC2 Alterations in Cancers. J Mol Biol 429, 1978–1993 (2017).

95. Piunti, A. & Shilatifard, A. The roles of Polycomb repressive complexes in mammalian development and cancer. Nature Reviews Molecular Cell Biology 22, 326–345 (2021).

96. Laugesen, A., Højfeldt, J. W. & Helin, K. Role of the Polycomb Repressive Complex 2 (PRC2) in Transcriptional Regulation and Cancer. Cold Spring Harb Perspect Med 6, (2016).

97. Visser, H. P. et al. The Polycomb group protein EZH2 is upregulated in proliferating, cultured human mantle cell lymphoma. Br J Haematol 112, 950–958 (2001).

98. van Kemenade, F. J. et al. Coexpression of BMI-1 and EZH2 polycomb-group proteins is associated with cycling cells and degree of malignancy in B-cell non-Hodgkin lymphoma. Blood 97, 3896–3901 (2001).

99. Varambally, S. et al. The polycomb group protein EZH2 is involved in progression of prostate cancer. Nature 419, 624–629 (2002).

100. Jiang, T. et al. Prognostic value of high EZH2 expression in patients with different types of cancer: a systematic review with meta-analysis. Oncotarget 7, 4584–4597 (2016).

101. Shi, J. et al. The Polycomb complex PRC2 supports aberrant self-renewal in a mouse model of MLL-AF9;Nras(G12D) acute myeloid leukemia. Oncogene 32, 930–938 (2013).

102. Danis, E. et al. Inactivation of Eed impedes MLL-AF9-mediated leukemogenesis through Cdkn2a-dependent and Cdkn2a-independent mechanisms in a murine model. Exp Hematol 43, 930-935.e6 (2015).

103. Pemov, A. et al. Low mutation burden and frequent loss of CDKN2A/B and SMARCA2, but not PRC2, define premalignant neurofibromatosis type 1-associated atypical neurofibromas. Neuro Oncol 21, 981–992 (2019).

104. Lee, W. et al. PRC2 is recurrently inactivated through EED or SUZ12 loss in malignant peripheral nerve sheath tumors. Nat Genet 46, 1227–1232 (2014).

105. Prieto-Granada, C. N. et al. Loss of H3K27me3 Expression Is a Highly Sensitive Marker for Sporadic and Radiation-induced MPNST. Am J Surg Pathol 40, 479–489 (2016).

106. Zhang, M. et al. Somatic mutations of SUZ12 in malignant peripheral nerve sheath tumors. Nat Genet 46, 1170–1172 (2014).

107. Mito, J. K., Qian, X., Doyle, L. A., Hornick, J. L. & Jo, V. Y. Role of Histone H3K27 Trimethylation Loss as a Marker for Malignant Peripheral Nerve Sheath Tumor in Fine-Needle Aspiration and Small Biopsy Specimens. Am J Clin Pathol 148, 179–189 (2017).

108. De Raedt, T. et al. PRC2 loss amplifies Ras-driven transcription and confers sensitivity to BRD4-based therapies. Nature 514, 247–251 (2014).

109. Krug, B., Harutyunyan, A. S., Deshmukh, S. & Jabado, N. Polycomb repressive complex 2 in the driver's seat of childhood and young adult brain tumours. Trends Cell Biol (2021).

110. Louis, D. N. et al. The 2016 World Health Organization Classification of Tumors of the Central Nervous System: a summary. Acta Neuropathol 131, 803–820 (2016).

111. Wu, G. et al. The genomic landscape of diffuse intrinsic pontine glioma and pediatric non-brainstem high-grade glioma. Nat Genet 46, 444–450 (2014).

112. Schwartzentruber, J. et al. Driver mutations in histone H3.3 and chromatin remodelling genes in paediatric glioblastoma. Nature 482, 226–231 (2012).

113. Dufour, C. et al. Identification of prognostic markers in diffuse midline gliomas H3K27M-mutant. Brain Pathol 30, 179–190 (2020).

114. Marchione, D. M. et al. Histone H3K27 dimethyl loss is highly specific for malignant peripheral nerve sheath tumor and distinguishes true PRC2 loss from isolated H3K27 trimethyl loss. Mod Pathol 32, 1434–1446 (2019).

115. Lewis, P. W. et al. Inhibition of PRC2 activity by a gain-of-function H3 mutation found in pediatric glioblastoma. Science 340, 857–861 (2013).

116. Bender, S. et al. Reduced H3K27me3 and DNA hypomethylation are major drivers of gene expression in K27M mutant pediatric high-grade gliomas. Cancer Cell 24, 660–672 (2013).

117. Funato, K., Major, T., Lewis, P. W., Allis, C. D. & Tabar, V. Use of human embryonic stem cells to model pediatric gliomas with H3.3K27M histone mutation. Science 346, 1529–1533 (2014).

118. Lee, C.-H. et al. Automethylation of PRC2 promotes H3K27 methylation and is impaired in H3K27M pediatric glioma. Genes Dev 33, 1428–1440 (2019).

119. Harutyunyan, A. S. et al. H3K27M induces defective chromatin spread of PRC2-mediated repressive H3K27me2/me3 and is essential for glioma tumorigenesis. Nat Commun 10, 1262 (2019).

120. Justin, N. et al. Structural basis of oncogenic histone H3K27M inhibition of human polycomb repressive complex 2. Nat Commun 7, 11316 (2016).

121. Herz, H.-M. et al. Histone H3 lysine-to-methionine mutants as a paradigm to study chromatin signaling. Science 345, 1065–1070 (2014).

122. Wang, X. et al. Molecular analysis of PRC2 recruitment to DNA in chromatin and its inhibition by RNA. Nat Struct Mol Biol 24, 1028–1038 (2017).

123. Mohammad, F. et al. EZH2 is a potential therapeutic target for H3K27M-mutant pediatric gliomas. Nat Med 23, 483–492 (2017).

124. Krug, B. et al. Pervasive H3K27 Acetylation Leads to ERV Expression and a Therapeutic Vulnerability in H3K27M Gliomas. Cancer Cell 35, 782-797.e8 (2019).

125. Brien, G. L. et al. Simultaneous disruption of PRC2 and enhancer function underlies histone H3.3-K27M oncogenic activity in human hindbrain neural stem cells. Nature Genetics 53, 1221–1232 (2021).

126. Piunti, A. et al. Therapeutic targeting of polycomb and BET bromodomain proteins in diffuse intrinsic pontine gliomas. Nat Med 23, 493–500 (2017).

127. Harutyunyan, A. S. et al. H3K27M in Gliomas Causes a One-Step Decrease in H3K27 Methylation and Reduced Spreading within the Constraints of H3K36 Methylation. Cell Rep 33, 108390 (2020).

128. Stafford, J. M. et al. Multiple modes of PRC2 inhibition elicit global chromatin alterations in H3K27M pediatric glioma. Sci Adv 4, eaau5935 (2018).

129. Lu, C. et al. Histone H3K36 mutations promote sarcomagenesis through altered histone methylation landscape. Science 352, 844–849 (2016).

130. Nikolaev, A., Fiveash, J. B. & Yang, E. S. Combined Targeting of Mutant p53 and Jumonji Family Histone Demethylase Augments Therapeutic Efficacy of Radiation in H3K27M DIPG. Int J Mol Sci 21, (2020).

131. Pajtler, K. W. et al. Molecular heterogeneity and CXorf67 alterations in posterior fossa group A (PFA) ependymomas. Acta Neuropathol 136, 211–226 (2018).

132. Jain, S. U. et al. PFA ependymoma-associated protein EZHIP inhibits PRC2 activity through a H3 K27M-like mechanism. Nat Commun 10, 2146 (2019).

133. Jain, S. U. et al. H3 K27M and EZHIP Impede H3K27-Methylation Spreading by Inhibiting Allosterically Stimulated PRC2. Mol Cell 80, 726-735.e7 (2020).

134. Piunti, A. et al. CATACOMB: An endogenous inducible gene that antagonizes H3K27 methylation activity of Polycomb repressive complex 2 via an H3K27M-like mechanism. Sci Adv 5, eaax2887 (2019).

135. Michealraj, K. A. et al. Metabolic Regulation of the Epigenome Drives Lethal Infantile Ependymoma. Cell 181, 1329-1345.e24 (2020).

136. Kaito, S. & Iwama, A. Pathogenic Impacts of Dysregulated Polycomb Repressive Complex Function in Hematological Malignancies. Int J Mol Sci 22, (2020).

137. Basheer, F. et al. Contrasting requirements during disease evolution identify EZH2 as a therapeutic target in AML. J Exp Med 216, 966–981 (2019).

138. Tanaka, S. et al. Ezh2 augments leukemogenicity by reinforcing differentiation blockage in acute myeloid leukemia. Blood 120, 1107–1117 (2012).

139. Neff, T. et al. Polycomb repressive complex 2 is required for MLL-AF9 leukemia. Proc Natl Acad Sci U S A 109, 5028–5033 (2012).

140. Khan, S. N. et al. Multiple mechanisms deregulate EZH2 and histone H3 lysine 27 epigenetic changes in myeloid malignancies. Leukemia 27, 1301–1309 (2013).

141. Venney, D., Mohd-Sarip, A. & Mills, K. I. The Impact of Epigenetic Modifications in Myeloid Malignancies. Int J Mol Sci 22, (2021).

142. Lindsley, R. C. et al. Acute myeloid leukemia ontogeny is defined by distinct somatic mutations. Blood 125, 1367–1376 (2015).

143. Muto, T. et al. Concurrent loss of Ezh2 and Tet2 cooperates in the pathogenesis of myelodysplastic disorders. J Exp Med 210, 2627–2639 (2013).

144. Sashida, G. et al. Ezh2 loss promotes development of myelodysplastic syndrome but attenuates its predisposition to leukaemic transformation. Nat Commun 5, 4177 (2014).

145. Oshima, M. et al. Ezh2 regulates the Lin28/let-7 pathway to restrict activation of fetal gene signature in adult hematopoietic stem cells. Exp Hematol 44, 282-296.e3 (2016).

146. Ntziachristos, P. et al. Genetic inactivation of the polycomb repressive complex 2 in T cell acute lymphoblastic leukemia. Nat Med 18, 298–301 (2012).

147. Wang, C. et al. Ezh2 loss propagates hypermethylation at T cell differentiation-regulating genes to promote leukemic transformation. J Clin Invest 128, 3872–3886 (2018).

148. Boileau, M. et al. Mutant H3 histones drive human pre-leukemic hematopoietic stem cell expansion and promote leukemic aggressiveness. Nat Commun 10, 2891 (2019).

149. Lehnertz, B. et al. H3K27M/I mutations promote context-dependent transformation in acute myeloid leukemia with RUNX1 alterations. Blood 130, 2204–2214 (2017).

150. Hassan, A., Pestana, R. C. & Parkes, A. Systemic Options for Malignant Peripheral Nerve Sheath Tumors. Curr Treat Options Oncol 22, 33 (2021).

151. Wojcik, J. B. et al. Epigenomic Reordering Induced by Polycomb Loss Drives Oncogenesis but Leads to Therapeutic Vulnerabilities in Malignant Peripheral Nerve Sheath Tumors. Cancer Res 79, 3205–3219 (2019).

152. Kochat, V. et al. Enhancer reprogramming in PRC2-deficient malignant peripheral nerve sheath tumors induces a targetable de-differentiated state. Acta Neuropathol (2021) doi:10.1007/s00401-021-02341-z.

153. Patel, A. J. et al. BET bromodomain inhibition triggers apoptosis of NF1-associated malignant peripheral nerve sheath tumors through Bim induction. Cell Rep 6, 81–92 (2014).

154. University of Texas Southwestern Medical Center. A Phase 2 Study of CPI-0610, a Small Molecule Inhibitor of Bromodomain and Extra-Terminal (BET) Proteins, in Patients With Malignant Peripheral Nerve Sheath Tumors. https://clinicaltrials.gov/ct2/show/NCT02986919 (2018).

155. Lopez, G. et al. Autophagic survival in resistance to histone deacetylase inhibitors: novel strategies to treat malignant peripheral nerve sheath tumors. Cancer Res 71, 185–196 (2011).

156. Brown, Z. Z. et al. Strategy for 'detoxification' of a cancer-derived histone mutant based on mapping its interaction with the methyltransferase PRC2. J Am Chem Soc 136, 13498–13501 (2014).

157. Nagaraja, S. et al. Transcriptional Dependencies in Diffuse Intrinsic Pontine Glioma. Cancer Cell 31, 635-652.e6 (2017).

158. Hashizume, R. et al. Pharmacologic inhibition of histone demethylation as a therapy for pediatric brainstem glioma. Nat Med 20, 1394–1396 (2014).

159. Wiese, M. et al. No Significant Cytotoxic Effect of the EZH2 Inhibitor Tazemetostat (EPZ-6438) on Pediatric Glioma Cells with Wildtype Histone 3 or Mutated Histone 3.3. Klin Padiatr 228, 113–117 (2016).

160. Ochs, K. et al. K27M-mutant histone-3 as a novel target for glioma immunotherapy. OncoImmunology 6, e1328340 (2017).

161. Chheda, Z. S. et al. Novel and shared neoantigen derived from histone 3 variant H3.3K27M mutation for glioma T cell therapy. Journal of Experimental Medicine 215, 141–157 (2017).

162. Majzner, R. G. et al. CAR T Cells Targeting B7-H3, a Pan-Cancer Antigen, Demonstrate Potent Preclinical Activity Against Pediatric Solid Tumors and Brain Tumors. Clin Cancer Res 25, 2560–2574 (2019).

163. Mount, C. W. et al. Potent antitumor efficacy of anti-GD2 CAR T cells in H3-K27M+ diffuse midline gliomas. Nat Med 24, 572–579 (2018).

164. Göllner, S. et al. Loss of the histone methyltransferase EZH2 induces resistance to multiple drugs in acute myeloid leukemia. Nat Med 23, 69–78 (2017).

165. Lambert, M. et al. Direct and Indirect Targeting of HOXA9 Transcription Factor in Acute Myeloid Leukemia. Cancers (Basel) 11, (2019).

166. Li, Y. et al. Therapeutic potential of GSK-J4, a histone demethylase KDM6B/JMJD3 inhibitor, for acute myeloid leukemia. J Cancer Res Clin Oncol 144, 1065–1077 (2018).

167. Ezponda, T. et al. UTX/KDM6A Loss Enhances the Malignant Phenotype of Multiple Myeloma and Sensitizes Cells to EZH2 inhibition. Cell Rep 21, 628–640 (2017).

168. Scott, M. T. et al. Epigenetic Reprogramming Sensitizes CML Stem Cells to Combined EZH2 and Tyrosine Kinase Inhibition. Cancer Discov 6, 1248–1257 (2016).

169. Gu, Z. et al. Loss of EZH2 Reprograms BCAA Metabolism to Drive Leukemic Transformation. Cancer Discov 9, 1228–1247 (2019).

170. Smith, A. et al. Lymphoma incidence, survival and prevalence 2004-2014: sub-type analyses from the UK's Haematological Malignancy Research Network. Br J Cancer 112, 1575–1584 (2015).

171. Le Guyader-Peyrou, S. et al. Cancer incidence in France over the 1980-2012 period: Hematological malignancies. Rev Epidemiol Sante Publique 64, 103–112 (2016).

172. Swerdlow, S. H. et al. WHO classification of tumours of haematopoietic and lymphoid tissues. (International Agency for Research on Cancer, 2017).

173. Follicular Lymphoma - Cancer Stat Facts. SEER https://seer.cancer.gov/statfacts/html/follicular.html.

174. Goldin, L. R., Björkholm, M., Kristinsson, S. Y., Turesson, I. & Landgren, O. Highly increased familial risks for specific lymphoma subtypes. Br J Haematol 146, 91–94 (2009).

175. Cerhan, J. R. & Slager, S. L. Familial predisposition and genetic risk factors for lymphoma. Blood 126, 2265–2273 (2015).

176. Skibola, C. F. et al. Genetic variants at 6p21.33 are associated with susceptibility to follicular lymphoma. Nat Genet 41, 873–875 (2009).

177. Wang, S. S. et al. HLA Class I and II Diversity Contributes to the Etiologic Heterogeneity of Non-Hodgkin Lymphoma Subtypes. Cancer Res 78, 4086–4096 (2018).

178. ’t Mannetje, A. et al. Occupation and Risk of Non-Hodgkin Lymphoma and Its Subtypes: A Pooled Analysis from the InterLymph Consortium. Environ Health Perspect 124, 396–405 (2016).

179. Koutros, S. et al. Non-Hodgkin lymphoma risk and organophosphate and carbamate insecticide use in the north American pooled project. Environ Int 127, 199–205 (2019).

180. Leon, M. E. et al. Pesticide use and risk of non-Hodgkin lymphoid malignancies in agricultural cohorts from France, Norway and the USA: a pooled analysis from the AGRICOH consortium. Int J Epidemiol 48, 1519–1535 (2019).

181. Moubadder, L., McCullough, L. E., Flowers, C. R. & Koff, J. L. Linking Environmental Exposures to Molecular Pathogenesis in Non-Hodgkin Lymphoma Subtypes. Cancer Epidemiol Biomarkers Prev 29, 1844–1855 (2020).

182. Linet, M. S. et al. Medical history, lifestyle, family history, and occupational risk factors for follicular lymphoma: the InterLymph Non-Hodgkin Lymphoma Subtypes Project. J Natl Cancer Inst Monogr 2014, 26–40 (2014).

183. Din, L. et al. Genetic overlap between autoimmune diseases and non-Hodgkin lymphoma subtypes. Genet Epidemiol 43, 844–863 (2019).

184. Chiu, B. C.-H. et al. Agricultural pesticide use and risk of t(14;18)-defined subtypes of non-Hodgkin lymphoma. Blood 108, 1363–1369 (2006).

185. Agopian, J. et al. Agricultural pesticide exposure and the molecular connection to lymphomagenesis. J Exp Med 206, 1473–1483 (2009).

186. Czuczman, M. S. et al. Treatment of patients with low-grade B-cell lymphoma with the combination of chimeric anti-CD20 monoclonal antibody and CHOP chemotherapy. J Clin Oncol 17, 268–276 (1999).

187. Juárez-Salcedo, L. M., Conde-Royo, D., Quiroz-Cervantes, K. & Dalia, S. Use of anti-CD20 therapy in follicular and marginal zone lymphoma: a review of the literature. Drugs Context 9, (2020).

188. Choi, S. M., Betz, B. L. & Perry, A. M. Follicular Lymphoma Diagnostic Caveats and Updates. Arch Pathol Lab Med 142, 1330–1340 (2018).

189. Federico, M. et al. Rituximab and the risk of transformation of follicular lymphoma: a retrospective pooled analysis. Lancet Haematol 5, e359–e367 (2018).

190. Payne, K. et al. BIOMED-2 PCR assays for IGK gene rearrangements are essential for B-cell clonality analysis in follicular lymphoma. Br J Haematol 155, 84–92 (2011).

191. Dreyling, M. et al. Newly diagnosed and relapsed follicular lymphoma: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. Ann Oncol 32, 298–308 (2021).

192. Meignan, M. et al. Baseline Metabolic Tumor Volume Predicts Outcome in High-Tumor-Burden Follicular Lymphoma: A Pooled Analysis of Three Multicenter Studies. J Clin Oncol 34, 3618–3626 (2016).

193. Batlevi, C. L. et al. Follicular lymphoma in the modern era: survival, treatment outcomes, and identification of high-risk subgroups. Blood Cancer Journal 10, 1–12 (2020).

194. Solal-Céligny, P. et al. Follicular lymphoma international prognostic index. Blood 104, 1258–1265 (2004).

195. Federico, M. et al. Follicular lymphoma international prognostic index 2: a new prognostic index for follicular lymphoma developed by the international follicular lymphoma prognostic factor project. J Clin Oncol 27, 4555–4562 (2009).

196. Pastore, A. et al. Integration of gene mutations in risk prognostication for patients receiving first-line immunochemotherapy for follicular lymphoma: a retrospective analysis of a prospective clinical trial and validation in a population-based registry. Lancet Oncol 16, 1111–1122 (2015).

197. Mondello, P. et al. Lack of intrafollicular memory CD4 + T cells is predictive of early clinical failure in newly diagnosed follicular lymphoma. Blood Cancer Journal 11, 1–11 (2021).

198. Bachy, E. et al. A simplified scoring system in de novo follicular lymphoma treated initially with immunochemotherapy. Blood 132, 49–58 (2018).

199. Huet, S. et al. A gene-expression profiling score for prediction of outcome in patients with follicular lymphoma: a retrospective training and validation analysis in three international cohorts. Lancet Oncol 19, 549–561 (2018).

200. Delfau-Larue, M.-H. et al. Total metabolic tumor volume, circulating tumor cells, cell-free DNA: distinct prognostic value in follicular lymphoma. Blood Adv 2, 807–816 (2018).

201. Jurinovic, V. et al. Clinicogenetic risk models predict early progression of follicular lymphoma after first-line immunochemotherapy. Blood 128, 1112–1120 (2016).

202. Hoskin, P. et al. 4 Gy versus 24 Gy radiotherapy for follicular and marginal zone lymphoma (FoRT): long-term follow-up of a multicentre, randomised, phase 3, non-inferiority trial. Lancet Oncol 22, 332–340 (2021).

203. Friedberg, J. W. et al. Effectiveness of first-line management strategies for stage I follicular lymphoma: analysis of the National LymphoCare Study. J Clin Oncol 30, 3368–3375 (2012).

204. Hiddemann, W. et al. Frontline therapy with rituximab added to the combination of cyclophosphamide, doxorubicin, vincristine, and prednisone (CHOP) significantly improves the outcome for patients with advanced-stage follicular lymphoma compared with therapy with CHOP alone: results of a prospective randomized study of the German Low-Grade Lymphoma Study Group. Blood 106, 3725–3732 (2005).

205. Hiddemann, W. et al. Immunochemotherapy With Obinutuzumab or Rituximab for Previously Untreated Follicular Lymphoma in the GALLIUM Study: Influence of Chemotherapy on Efficacy and Safety. J Clin Oncol 36, 2395–2404 (2018).

206. Rummel, M. J. et al. Bendamustine plus rituximab versus CHOP plus rituximab as first-line treatment for patients with indolent and mantle-cell lymphomas: an open-label, multicentre, randomised, phase 3 non-inferiority trial. Lancet 381, 1203–1210 (2013).

207. Morschhauser, F. et al. Rituximab plus Lenalidomide in Advanced Untreated Follicular Lymphoma. N Engl J Med 379, 934–947 (2018).

208. Bachy, E. et al. Sustained Progression-Free Survival Benefit of Rituximab Maintenance in Patients With Follicular Lymphoma: Long-Term Results of the PRIMA Study. J Clin Oncol 37, 2815–2824 (2019).

209. Hirt, C. et al. Rituximab Maintenance Versus Observation After Immunochemotherapy (R-CHOP, R-MCP, and R-FCM) in Untreated Follicular Lymphoma Patients: A Randomized Trial of the Ostdeutsche Studiengruppe Hämatologie und Onkologie and the German Low-Grade Lymphoma Study Group. Hemasphere 5, e600 (2021).

210. Cheson, B. D. et al. Overall Survival Benefit in Patients With Rituximab-Refractory Indolent Non-Hodgkin Lymphoma Who Received Obinutuzumab Plus Bendamustine Induction and Obinutuzumab Maintenance in the GADOLIN Study. J Clin Oncol 36, 2259–2266 (2018).

211. Gopal, A. K. et al. PI3Kδ inhibition by idelalisib in patients with relapsed indolent lymphoma. N Engl J Med 370, 1008–1018 (2014).

212. Dreyling, M. et al. Long-term safety and efficacy of the PI3K inhibitor copanlisib in patients with relapsed or refractory indolent lymphoma: 2-year follow-up of the CHRONOS-1 study. Am J Hematol 95, 362–371 (2020).

213. Flinn, I. W. et al. DYNAMO: A Phase II Study of Duvelisib (IPI-145) in Patients With Refractory Indolent Non-Hodgkin Lymphoma. J Clin Oncol 37, 912–922 (2019).

214. Casulo, C. et al. Autologous Transplantation in Follicular Lymphoma with Early Therapy Failure: A National LymphoCare Study and Center for International Blood and Marrow Transplant Research Analysis. Biol Blood Marrow Transplant 24, 1163–1171 (2018).

215. Leonard, J. P. et al. AUGMENT: A Phase III Study of Lenalidomide Plus Rituximab Versus Placebo Plus Rituximab in Relapsed or Refractory Indolent Lymphoma. J Clin Oncol 37, 1188–1199 (2019).

216. Lampson, B. L. & Brown, J. R. PI3Kδ-selective and PI3Kα/δ-combinatorial inhibitors in clinical development for B-cell non-Hodgkin lymphoma. Expert Opin Investig Drugs 26, 1267–1279 (2017).

217. Gong, J., Chehrazi-Raffle, A., Reddi, S. & Salgia, R. Development of PD-1 and PD-L1 inhibitors as a form of cancer immunotherapy: a comprehensive review of registration trials and future considerations. J Immunother Cancer 6, 8 (2018).

218. Advani, R. et al. CD47 Blockade by Hu5F9-G4 and Rituximab in Non-Hodgkin's Lymphoma. N Engl J Med 379, 1711–1721 (2018).

219. Schuster, S. J. Bispecific antibodies for the treatment of lymphomas: Promises and challenges. Hematol Oncol 39 Suppl 1, 113–116 (2021).

220. Schuster, S. J. et al. Chimeric Antigen Receptor T Cells in Refractory B-Cell Lymphomas. N Engl J Med 377, 2545–2554 (2017).

221. Ribrag, V. et al. Safety and efficacy of abexinostat, a pan-histone deacetylase inhibitor, in non-Hodgkin lymphoma and chronic lymphocytic leukemia: results of a phase II study. Haematologica 102, 903–909 (2017).

222. Huet, S., Sujobert, P. & Salles, G. From genetics to the clinic: a translational perspective on follicular lymphoma. Nat Rev Cancer 18, 224–239 (2018).

223. Lackraj, T., Goswami, R. & Kridel, R. Pathogenesis of follicular lymphoma. Best Pract Res Clin Haematol 31, 2–14 (2018).

224. Carbone, A. et al. Follicular lymphoma. Nat Rev Dis Primers 5, 83 (2019).

225. Tsujimoto, Y., Finger, L. R., Yunis, J., Nowell, P. C. & Croce, C. M. Cloning of the chromosome breakpoint of neoplastic B cells with the t(14;18) chromosome translocation. Science 226, 1097–1099 (1984).

226. Tsujimoto, Y., Cossman, J., Jaffe, E. & Croce, C. M. Involvement of the bcl-2 gene in human follicular lymphoma. Science 228, 1440–1443 (1985).

227. Tsujimoto, Y., Gorham, J., Cossman, J., Jaffe, E. & Croce, C. M. The t(14;18) chromosome translocations involved in B-cell neoplasms result from mistakes in VDJ joining. Science 229, 1390–1393 (1985).

228. Llambi, F. et al. A unified model of mammalian BCL-2 protein family interactions at the mitochondria. Mol Cell 44, 517–531 (2011).

229. Limpens, J. et al. Lymphoma-associated translocation t(14;18) in blood B cells of normal individuals. Blood 85, 2528–2536 (1995).

230. Brisou, G., Nadel, B. & Roulland, S. The Premalignant Ancestor Cell of t(14;18)+ Lymphoma. Hemasphere 5, e579 (2021).

231. Roulland, S. et al. Follicular lymphoma-like B cells in healthy individuals: a novel intermediate step in early lymphomagenesis. J Exp Med 203, 2425–2431 (2006).

232. Hirt, C. et al. Risk of follicular lymphoma associated with BCL2 translocations in peripheral blood. Leuk Lymphoma 56, 2625–2629 (2015).

233. Roulland, S. et al. t(14;18) Translocation: A predictive blood biomarker for follicular lymphoma. J Clin Oncol 32, 1347–1355 (2014).

234. Sungalee, S. et al. Germinal center reentries of BCL2-overexpressing B cells drive follicular lymphoma progression. J Clin Invest 124, 5337–5351 (2014).

235. Weigert, O. et al. Molecular ontogeny of donor-derived follicular lymphomas occurring after hematopoietic cell transplantation. Cancer Discov 2, 47–55 (2012).

236. Green, M. R. et al. Hierarchy in somatic mutations arising during genomic evolution and progression of follicular lymphoma. Blood 121, 1604–1611 (2013).

237. Green, M. R. et al. Mutations in early follicular lymphoma progenitors are associated with suppressed antigen presentation. Proc Natl Acad Sci U S A 112, E1116-1125 (2015).

238. Horton, S. J. et al. Early loss of Crebbp confers malignant stem cell properties on lymphoid progenitors. Nat Cell Biol 19, 1093–1104 (2017).

239. Horsman, D. E., Connors, J. M., Pantzar, T. & Gascoyne, R. D. Analysis of secondary chromosomal alterations in 165 cases of follicular lymphoma with t(14;18). Genes Chromosomes Cancer 30, 375–382 (2001).

240. Kridel, R. et al. Histological Transformation and Progression in Follicular Lymphoma: A Clonal Evolution Study. PLoS Med 13, (2016).

241. 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).

242. 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).

243. Morin, R. D. et al. Frequent mutation of histone-modifying genes in non-Hodgkin lymphoma. Nature 476, 298–303 (2011).

244. Lunning, M. A. & Green, M. R. Mutation of chromatin modifiers; an emerging hallmark of germinal center B-cell lymphomas. Blood Cancer J 5, e361 (2015).

245. Froimchuk, E., Jang, Y. & Ge, K. Histone H3 lysine 4 methyltransferase KMT2D. Gene 627, 337–342 (2017).

246. Ford, D. J. & Dingwall, A. K. The cancer COMPASS: navigating the functions of MLL complexes in cancer. Cancer Genet 208, 178–191 (2015).

247. Zhang, J. et al. Disruption of KMT2D perturbs germinal center B cell development and promotes lymphomagenesis. Nat Med 21, 1190–1198 (2015).

248. 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).

249. García-Ramírez, I. et al. Crebbp loss cooperates with Bcl2 overexpression to promote lymphoma in mice. Blood 129, 2645–2656 (2017).

250. Pasqualucci, L. et al. Inactivating mutations of acetyltransferase genes in B-cell lymphoma. Nature 471, 189–195 (2011).

251. Krysiak, K. et al. Recurrent somatic mutations affecting B-cell receptor signaling pathway genes in follicular lymphoma. Blood 129, 473–483 (2017).

252. Gounder, M. et al. Tazemetostat in advanced epithelioid sarcoma with loss of INI1/SMARCB1: an international, open-label, phase 2 basket study. Lancet Oncol 21, 1423–1432 (2020).

253. Kretzmer, H. et al. DNA-methylome analysis in Burkitt and follicular lymphomas identifies differentially methylated regions linked to somatic mutation and transcriptional control. Nat Genet 47, 1316–1325 (2015).

254. Hayslip, J. & Montero, A. Tumor suppressor gene methylation in follicular lymphoma: a comprehensive review. Molecular Cancer 5, 44 (2006).

255. De, S. et al. Aberration in DNA Methylation in B-Cell Lymphomas Has a Complex Origin and Increases with Disease Severity. PLOS Genetics 9, e1003137 (2013).

256. Basso, K. & Dalla-Favera, R. Roles of BCL6 in normal and transformed germinal center B cells. Immunol Rev 247, 172–183 (2012).

257. Ying, C. Y. et al. MEF2B mutations lead to deregulated expression of the oncogene BCL6 in diffuse large B cell lymphoma. Nat Immunol 14, 1084–1092 (2013).

258. Valls, E. et al. BCL6 Antagonizes NOTCH2 to Maintain Survival of Human Follicular Lymphoma Cells. Cancer Discov 7, 506–521 (2017).

259. Béguelin, W. et al. EZH2 and BCL6 Cooperate to Assemble CBX8-BCOR Complex to Repress Bivalent Promoters, Mediate Germinal Center Formation and Lymphomagenesis. Cancer Cell 30, 197–213 (2016).

260. Okosun, J. et al. Recurrent mTORC1-activating RRAGC mutations in follicular lymphoma. Nat Genet 48, 183–188 (2016).

261. Oricchio, E. et al. Genetic and epigenetic inactivation of SESTRIN1 controls mTORC1 and response to EZH2 inhibition in follicular lymphoma. Sci Transl Med 9, (2017).

262. Cheung, K.-J. J. et al. Acquired TNFRSF14 mutations in follicular lymphoma are associated with worse prognosis. Cancer Res 70, 9166–9174 (2010).

263. Cai, G. & Freeman, G. J. The CD160, BTLA, LIGHT/HVEM pathway: a bidirectional switch regulating T-cell activation. Immunol Rev 229, 244–258 (2009).

264. Murphy, K. M., Nelson, C. A. & Sedý, J. R. Balancing co-stimulation and inhibition with BTLA and HVEM. Nat Rev Immunol 6, 671–681 (2006).

265. Boice, M. et al. Loss of the HVEM Tumor Suppressor in Lymphoma and Restoration by Modified CAR-T Cells. Cell 167, 405-418.e13 (2016).

266. Wrench, D. et al. TNFRSF14 and EZH2 Mutations, Chr2p Gain and Copy Number Changes Targeting Genes Whose Proteins Interact with the Microenvironment In Transformed Follicular Lymphoma. Blood 116, 799 (2010).

267. Yildiz, M. et al. Activating STAT6 mutations in follicular lymphoma. Blood 125, 668–679 (2015).

268. Mottok, A. et al. Inactivating SOCS1 mutations are caused by aberrant somatic hypermutation and restricted to a subset of B-cell lymphoma entities. Blood 114, 4503–4506 (2009).

269. Li, H. et al. Mutations in linker histone genes HIST1H1 B, C, D, and E; OCT2 (POU2F2); IRF8; and ARID1A underlying the pathogenesis of follicular lymphoma. Blood 123, 1487–1498 (2014).

270. 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).

271. Quizon, N., Kwak, K., Shah, S., Singh, A. & Pierce, S. Bioengineered organoid models of human germinal centers. The Journal of Immunology 200, 120.14-120.14 (2018).

272. Dave, S. S. et al. Prediction of survival in follicular lymphoma based on molecular features of tumor-infiltrating immune cells. N Engl J Med 351, 2159–2169 (2004).

273. Harjunpää, A. et al. Differential gene expression in non-malignant tumour microenvironment is associated with outcome in follicular lymphoma patients treated with rituximab and CHOP. Br J Haematol 135, 33–42 (2006).

274. Yang, Z.-Z., Novak, A. J., Ziesmer, S. C., Witzig, T. E. & Ansell, S. M. Malignant B Cells Skew the Balance of Regulatory T Cells and TH17 Cells in B-Cell Non-Hodgkin's Lymphoma. Cancer Res 69, 5522–5530 (2009).

275. Amé-Thomas, P. & Tarte, K. The yin and the yang of follicular lymphoma cell niches: role of microenvironment heterogeneity and plasticity. Semin Cancer Biol 24, 23–32 (2014).

276. de Jong, D. & Fest, T. The microenvironment in follicular lymphoma. Best Pract Res Clin Haematol 24, 135–146 (2011).

277. Pangault, C. et al. Integrative Analysis of Cell Crosstalk within Follicular Lymphoma Cell Niche: Towards a Definition of the FL Supportive Synapse. Cancers (Basel) 12, (2020).

278. Pandey, S. et al. IL-4/CXCL12 loop is a key regulator of lymphoid stroma function in follicular lymphoma. Blood 129, 2507–2518 (2017).

279. Carbone, A., Gloghini, A., Gruss, H. J. & Pinto, A. CD40 ligand is constitutively expressed in a subset of T cell lymphomas and on the microenvironmental reactive T cells of follicular lymphomas and Hodgkin's disease. Am J Pathol 147, 912–922 (1995).

280. Radcliffe, C. M. et al. Human follicular lymphoma cells contain oligomannose glycans in the antigen-binding site of the B-cell receptor. J Biol Chem 282, 7405–7415 (2007).

281. Coelho, V. et al. Glycosylation of surface Ig creates a functional bridge between human follicular lymphoma and microenvironmental lectins. Proc Natl Acad Sci U S A 107, 18587–18592 (2010).

282. Linley, A. et al. Lectin binding to surface Ig variable regions provides a universal persistent activating signal for follicular lymphoma cells. Blood 126, 1902–1910 (2015).

283. Devan, J., Janikova, A. & Mraz, M. New concepts in follicular lymphoma biology: From BCL2 to epigenetic regulators and non-coding RNAs. Semin Oncol 45, 291–302 (2018).

284. Pasqualucci, L. et al. Genetics of follicular lymphoma transformation. Cell Rep 6, 130–140 (2014).

285. González-Rincón, J. et al. Unraveling transformation of follicular lymphoma to diffuse large B-cell lymphoma. PLoS One 14, e0212813 (2019).

286. Mochizuki-Kashio, M. et al. Ezh2 loss in hematopoietic stem cells predisposes mice to develop heterogeneous malignancies in an Ezh1-dependent manner. Blood 126, 1172–1183 (2015).

287. Kamminga, L. M. et al. The Polycomb group gene Ezh2 prevents hematopoietic stem cell exhaustion. Blood 107, 2170–2179 (2006).

288. Majewski, I. J. et al. Opposing roles of polycomb repressive complexes in hematopoietic stem and progenitor cells. Blood 116, 731–739 (2010).

289. Yoshida, H. et al. The cis-Regulatory Atlas of the Mouse Immune System. Cell 176, 897-912.e20 (2019).

290. Lee, S. C. W. et al. Polycomb repressive complex 2 component Suz12 is required for hematopoietic stem cell function and lymphopoiesis. Blood 126, 167–175 (2015).

291. Ebert, A., Hill, L. & Busslinger, M. Spatial Regulation of V-(D)J Recombination at Antigen Receptor Loci. Adv Immunol 128, 93–121 (2015).

292. Jacobsen, J. A. et al. EZH2 Regulates the Developmental Timing of Effectors of the Pre-Antigen Receptor Checkpoints. J Immunol 198, 4682–4691 (2017).

293. Mandal, M. et al. Epigenetic repression of the Igk locus by STAT5-mediated recruitment of the histone methyltransferase Ezh2. Nat Immunol 12, 1212–1220 (2011).

294. 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).

295. Caganova, M. et al. Germinal center dysregulation by histone methyltransferase EZH2 promotes lymphomagenesis. J Clin Invest 123, 5009–5022 (2013).

296. Béguelin, W. et al. EZH2 is required for germinal center formation and somatic EZH2 mutations promote lymphoid transformation. Cancer Cell 23, 677–692 (2013).

297. Béguelin, W. et al. EZH2 enables germinal centre formation through epigenetic silencing of CDKN1A and an Rb-E2F1 feedback loop. Nat Commun 8, 877 (2017).

298. Herviou, L., Jourdan, M., Martinez, A.-M., Cavalli, G. & Moreaux, J. EZH2 is overexpressed in transitional preplasmablasts and is involved in human plasma cell differentiation. Leukemia 33, 2047–2060 (2019).

299. Guo, M. et al. EZH2 Represses the B Cell Transcriptional Program and Regulates Antibody-Secreting Cell Metabolism and Antibody Production. J Immunol 200, 1039–1052 (2018).

300. Nutt, S. L., Keenan, C., Chopin, M. & Allan, R. S. EZH2 function in immune cell development. Biol Chem 401, 933–943 (2020).

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

302. Huet, S. et al. EZH2 alterations in follicular lymphoma: biological and clinical correlations. Blood Cancer J 7, e555 (2017).

303. Rada, C., Di Noia, J. M. & Neuberger, M. S. Mismatch recognition and uracil excision provide complementary paths to both Ig switching and the A/T-focused phase of somatic mutation. Mol Cell 16, 163–171 (2004).

304. Bouska, A. et al. Combined copy number and mutation analysis identifies oncogenic pathways associated with transformation of follicular lymphoma. Leukemia 31, 83–91 (2017).

305. Jones, R. S. & Gelbart, W. M. Genetic Analysis of the Enhancer of Zeste Locus and Its Role in Gene Regulation in Drosophila Melanogaster. Genetics 126, 185–199 (1990).

306. Joshi, P. et al. Dominant alleles identify SET domain residues required for histone methyltransferase of Polycomb repressive complex 2. J Biol Chem 283, 27757–27766 (2008).

307. McCabe, M. T. et al. Mutation of A677 in histone methyltransferase EZH2 in human B-cell lymphoma promotes hypertrimethylation of histone H3 on lysine 27 (H3K27). Proc Natl Acad Sci U S A 109, 2989–2994 (2012).

308. Wu, H. et al. Structure of the catalytic domain of EZH2 reveals conformational plasticity in cofactor and substrate binding sites and explains oncogenic mutations. PLoS One 8, e83737 (2013).

309. Antonysamy, S. et al. Structural context of disease-associated mutations and putative mechanism of autoinhibition revealed by X-ray crystallographic analysis of the EZH2-SET domain. PLoS One 8, e84147 (2013).

310. Brooun, A. et al. Polycomb repressive complex 2 structure with inhibitor reveals a mechanism of activation and drug resistance. Nat Commun 7, 11384 (2016).

311. Sahasrabuddhe, A. A. et al. Oncogenic Y641 mutations in EZH2 prevent Jak2/β-TrCP-mediated degradation. Oncogene 34, 445–454 (2015).

312. Pasini, D. et al. Characterization of an antagonistic switch between histone H3 lysine 27 methylation and acetylation in the transcriptional regulation of Polycomb group target genes. Nucleic Acids Res 38, 4958–4969 (2010).

313. Araf, S. et al. Genomic profiling reveals spatial intra-tumor heterogeneity in follicular lymphoma. Leukemia 32, 1261–1265 (2018).

314. Bödör, C. et al. EZH2 Y641 mutations in follicular lymphoma. Leukemia 25, 726–729 (2011).

315. Béguelin, W. et al. Mutant EZH2 Induces a Pre-malignant Lymphoma Niche by Reprogramming the Immune Response. Cancer Cell 37, 655-673.e11 (2020).

316. Velichutina, I. et al. EZH2-mediated epigenetic silencing in germinal center B cells contributes to proliferation and lymphomagenesis. Blood 116, 5247–5255 (2010).

317. Berg, T. et al. A transgenic mouse model demonstrating the oncogenic role of mutations in the polycomb-group gene EZH2 in lymphomagenesis. Blood 123, 3914–3924 (2014).

318. Lee, S. C. W. et al. Polycomb repressive complex 2 (PRC2) suppresses Eμ-myc lymphoma. Blood 122, 2654–2663 (2013).

319. Calado, D. P. et al. The cell-cycle regulator c-Myc is essential for the formation and maintenance of germinal centers. Nat Immunol 13, 1092–1100 (2012).

320. Dominguez-Sola, D. et al. The proto-oncogene MYC is required for selection in the germinal center and cyclic reentry. Nat Immunol 13, 1083–1091 (2012).

321. Shaffer, A. L. et al. Signatures of the immune response. Immunity 15, 375–385 (2001).

322. Donaldson-Collier, M. C. et al. EZH2 oncogenic mutations drive epigenetic, transcriptional, and structural changes within chromatin domains. Nat Genet 51, 517–528 (2019).

323. Ennishi, D. et al. Molecular and Genetic Characterization of MHC Deficiency Identifies EZH2 as Therapeutic Target for Enhancing Immune Recognition. Cancer Discov 9, 546–563 (2019).

324. O'Riain, C. et al. Array-based DNA methylation profiling in follicular lymphoma. Leukemia 23, 1858–1866 (2009).

325. Tiffen, J., Wilson, S., Gallagher, S. J., Hersey, P. & Filipp, F. V. Somatic Copy Number Amplification and Hyperactivating Somatic Mutations of EZH2 Correlate With DNA Methylation and Drive Epigenetic Silencing of Genes Involved in Tumor Suppression and Immune Responses in Melanoma. Neoplasia 18, 121–132 (2016).

326. Hu, Y. et al. Integrated Whole-Exome and Transcriptome Sequencing of Sporadic Parathyroid Adenoma. Front Endocrinol (Lausanne) 12, 631680 (2021).

327. Fan, T. et al. EZH2-Dependent Suppression of a Cellular Senescence Phenotype in Melanoma Cells by Inhibition of p21/CDKN1A Expression. Mol Cancer Res 9, 418–429 (2011).

328. Harms, P. W. et al. Activating mutations of the oncogene EZH2 in cutaneous melanoma revealed by next generation sequencing. Human Pathology: Case Reports 1, 21–28 (2014).

329. Green, M. R. Chromatin modifying gene mutations in follicular lymphoma. Blood 131, 595–604 (2018).

330. Glazer, R. I. et al. 3-Deazaneplanocin: a new and potent inhibitor of S-adenosylhomocysteine hydrolase and its effects on human promyelocytic leukemia cell line HL-60. Biochem Biophys Res Commun 135, 688–694 (1986).

331. Tan, J. et al. Pharmacologic disruption of Polycomb-repressive complex 2-mediated gene repression selectively induces apoptosis in cancer cells. Genes Dev 21, 1050–1063 (2007).

332. Miranda, T. B. et al. DZNep is a global histone methylation inhibitor that reactivates developmental genes not silenced by DNA methylation. Mol Cancer Ther 8, 1579–1588 (2009).

333. Konze, K. D. et al. An orally bioavailable chemical probe of the Lysine Methyltransferases EZH2 and EZH1. ACS Chem Biol 8, 1324–1334 (2013).

334. Knutson, S. K. et al. A selective inhibitor of EZH2 blocks H3K27 methylation and kills mutant lymphoma cells. Nat Chem Biol 8, 890–896 (2012).

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

336. Qi, W. et al. Selective inhibition of Ezh2 by a small molecule inhibitor blocks tumor cells proliferation. Proc Natl Acad Sci U S A 109, 21360–21365 (2012).

337. Knutson, S. K. et al. Durable tumor regression in genetically altered malignant rhabdoid tumors by inhibition of methyltransferase EZH2. Proc Natl Acad Sci U S A 110, 7922–7927 (2013).

338. Knutson, S. K. et al. Selective inhibition of EZH2 by EPZ-6438 leads to potent antitumor activity in EZH2-mutant non-Hodgkin lymphoma. Mol Cancer Ther 13, 842–854 (2014).

339. Brach, D. et al. EZH2 Inhibition by Tazemetostat Results in Altered Dependency on B-cell Activation Signaling in DLBCL. Mol Cancer Ther 16, 2586–2597 (2017).

340. Bradley, W. D. et al. EZH2 inhibitor efficacy in non-Hodgkin's lymphoma does not require suppression of H3K27 monomethylation. Chem Biol 21, 1463–1475 (2014).

341. Gehling, V. S. et al. Discovery, design, and synthesis of indole-based EZH2 inhibitors. Bioorg Med Chem Lett 25, 3644–3649 (2015).

342. Garapaty-Rao, S. et al. Identification of EZH2 and EZH1 small molecule inhibitors with selective impact on diffuse large B cell lymphoma cell growth. Chem Biol 20, 1329–1339 (2013).

343. Kung, P.-P. et al. Optimization of Orally Bioavailable Enhancer of Zeste Homolog 2 (EZH2) Inhibitors Using Ligand and Property-Based Design Strategies: Identification of Development Candidate (R)-5,8-Dichloro-7-(methoxy(oxetan-3-yl)methyl)-2-((4-methoxy-6-methyl-2-oxo-1,2-dihydropyridin-3-yl)methyl)-3,4-dihydroisoquinolin-1(2H)-one (PF-06821497). J Med Chem 61, 650–665 (2018).

344. 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. Lancet Oncol 19, 649–659 (2018).

345. Morschhauser, F. et al. Tazemetostat for patients with relapsed or refractory follicular lymphoma: an open-label, single-arm, multicentre, phase 2 trial. Lancet Oncol 21, 1433–1442 (2020).

346. Sarkozy, C. et al. A LYSA Phase Ib Study of Tazemetostat (EPZ-6438) plus R-CHOP in Patients with Newly Diagnosed Diffuse Large B-Cell Lymphoma (DLBCL) with Poor Prognosis Features. Clin Cancer Res 26, 3145–3153 (2020).

347. McDonald, A. et al. Updated Report on Identification of Molecular Predictors of Tazemetostat Response in an Ongoing NHL Phase 2 Study. Blood 132, 4097–4097 (2018).

348. Yap, T. A. et al. Phase I Study of the Novel Enhancer of Zeste Homolog 2 (EZH2) Inhibitor GSK2816126 in Patients with Advanced Hematologic and Solid Tumors. Clin Cancer Res 25, 7331–7339 (2019).

349. Park, S. et al. Combination Treatment with GSK126 and Pomalidomide Induces B-Cell Differentiation in EZH2 Gain-of-Function Mutant Diffuse Large B-Cell Lymphoma. Cancers (Basel) 12, (2020).

350. Bisserier, M. & Wajapeyee, N. Mechanisms of resistance to EZH2 inhibitors in diffuse large B-cell lymphomas. Blood 131, 2125–2137 (2018).

351. Tong, K. I. et al. Combined EZH2 inhibition and Ikaros degradation leads to enhanced anti-tumor activity in diffuse large B-cell lymphoma. Clin Cancer Res (2021).

352. Yamagishi, M. et al. Targeting Excessive EZH1 and EZH2 Activities for Abnormal Histone Methylation and Transcription Network in Malignant Lymphomas. Cell Rep 29, 2321-2337.e7 (2019).

353. Jiang, Y. et al. CREBBP Inactivation Promotes the Development of HDAC3-Dependent Lymphomas. Cancer Discov 7, 38–53 (2017).

APPENDIX

EZH1/2 function mostly within canonical PRC2 and exhibit proliferation-dependent redundancy that shapes mutational signatures in cancer

Michel Wassef[a,b,1], Armelle Luscan[a,b,1], Setareh Aflaki[a,b], Dina Zielinski[a,b,c], Pascal W. T. C. Jansen[d], H. Irem Baymaz[d], Aude Battistella[a,b], Carole Kersouani[a,b], Nicolas Servant[a,c], Margaret R. Wallace[e], Pierre Romero[a,b], Olivier Kosmider[f], Pierre-Alexandre Just[g,h], Mikaël Hivelin[i,j], Sébastien Jacques[k], Anne Vincent-Salomon[a,b], Michiel Vermeulen[d], Michel Vidaud[i,l], Eric Pasmant[i,l,2], and Raphaël Margueron[a,b,2]

[a]Institut Curie, Paris Sciences et Lettres Research University, 75005 Paris, France; [b]INSERM U934/CNRS UMR3215, 75248 Paris, France; [c]INSERM U900, Mines ParisTech, 75248 Paris, France; [d]Department of Molecular Biology, Faculty of Science, Radboud Institute for Molecular Life Sciences, Oncode Institute, Radboud University Nijmegen, 6525 GA Nijmegen, The Netherlands; [e]Department of Molecular Genetics and Microbiology, University of Florida Genetics Institute, University of Florida Health Cancer Center, University of Florida, Gainesville, FL 32610; [f]Institut Cochin, Department Development, Reproduction and Cancer, and Service d'Hématologie Biologique, Hôpitaux Universitaires Paris Centre–Cochin, Assistance Publique–Hôpitaux de Paris, 75014 Paris, France; [g]Department of Pathology, Cochin Hospital, Hôpitaux Universitaires Paris Centre, Assistance Publique–Hôpitaux de Paris, 75014 Paris, France; [h]Faculty of Medicine, Paris Descartes University, 75006 Paris, France; [i]Institut Cochin, INSERM U1016, Paris Descartes University, 75014 Paris, France; [j]Department of Plastic, Reconstructive, and Aesthetic Surgery, Hôpital Européen Georges Pompidou, Assistance Publique–Hôpitaux de Paris, 75015 Paris, France; [k]Genomic Platform, Institut Cochin, INSERM U1016, CNRS UMR8104, Paris Descartes University, 75014 Paris, France; and [l]Department of Molecular Genetics, Cochin Hospital, Hôpitaux Universitaires Paris Centre, Assistance Publique–Hôpitaux de Paris, 75014 Paris, France

Edited by Danny Reinberg, New York University School of Medicine, New York, NY, and approved February 20, 2019 (received for review August 24, 2018)

Genetic mutations affecting chromatin modifiers are widespread in cancers. In malignant peripheral nerve sheath tumors (MPNSTs), Polycomb repressive complex 2 (PRC2), which plays a crucial role in gene silencing, is inactivated through recurrent mutations in core subunits embryonic ectoderm development (EED) and suppressor of zeste 12 homolog (SUZ12), but mutations in PRC2's main catalytic subunit enhancer of zeste homolog 2 (EZH2) have never been found. This is in contrast to myeloid and lymphoid malignancies, which harbor frequent loss-of-function mutations in EZH2. Here, we investigated whether the absence of EZH2 mutations in MPNST is due to a PRC2-independent (i.e., noncanonical) function of the enzyme or to redundancy with EZH1. We show that, in the absence of SUZ12, EZH2 remains bound to EED but loses its interaction with all other core and accessory PRC2 subunits. Through genetic and pharmacological analyses, we unambiguously establish that EZH2 is functionally inert in this context, thereby excluding a PRC2-independent function. Instead, we show that EZH1 and EZH2 are functionally redundant in the slowly proliferating MPNST precursors. We provide evidence that the compensatory function of EZH1 is alleviated upon higher proliferation. This work reveals how context-dependent redundancies can shape tumor-type specific mutation patterns in chromatin regulators.

chromatin | cancer | Polycomb | EZH2

It is estimated that over 25% of the most frequently mutated genes in cancers encode chromatin regulators (1, 2). While some of these mutations are recurrent in a wide range of cancers, many others are found in specific tumor types, suggesting a context-dependent function. This is well illustrated in the case of Polycomb repressive complex 2 (PRC2), a chromatin-modifying complex involved in maintaining transcriptional repression. PRC2 contains several essential subunits: embryonic ectoderm development (EED), suppressor of zeste 12 homolog (SUZ12), retinoblastoma-binding protein 4/7 (RBBP4/7), and two paralogous enzymatic subunits enhancer of zeste homolog 1 and 2 (EZH1 and EZH2). EZH1 and EZH2 assemble into alternative PRC2 complexes with similar composition that both catalyze methylation of lysine 27 on histone 3 (H3K27) (3, 4). In addition, several accessory subunits associate with PRC2, assist in its recruitment, and/or modulate its enzymatic activity (5). Methylation of H3K27 is essential for Polycomb-mediated silencing (6). Alterations of the PRC2 complex have been reported across different malignancies, and each alteration displays striking tumor type specificity (reviewed in ref. 7). Gain-of-function mutations in EZH2 have been reported in follicular lymphoma, diffuse large B cell lymphoma, and a small subset of melanoma, while functionally similar mutations affecting EZH1 have been found in autonomous thyroid adenomas. In contrast, recurrent loss-of-function mutations in PRC2 genes occur in myeloid malignancies, T cell acute lymphoblastic leukemia (T-ALL), and malignant peripheral nerve sheath tumors (MPNSTs) (7).

MPNSTs are aggressive soft-tissue sarcomas that develop either sporadically or from preexisting benign tumors called plexiform neurofibromas in patients with neurofibromatosis type 1 (NF1) (Online Mendelian Inheritance in Man 162200). Previous studies have identified recurrent biallelic mutations in *EED* and *SUZ12* that result in a complete loss of H3K27me3 in over 50% of MPNSTs (8–10). Loss of PRC2 function as well as co-occurring

Significance

It is proposed that chromatin modifiers can regulate transcription through different mechanisms sometimes referred as "canonical" (toward chromatin) and "noncanonical" (toward other proteins). However, their relative contribution to the overall function of a given chromatin modifier is often enigmatic. We focused on the Polycomb complex PRC2 to investigate this question. Our results indicate that the canonical activity of PRC2 is largely predominant, if not exclusive, and that the particular pattern of PRC2 mutations in cancer is due to proliferation-dependent redundancy between the two enzymatic subunits of the complex.

Author contributions: M.W., A.L., M. Vidaud, E.P., and R.M. designed research; M.W., A.L., S.A., P.W.T.C.J., H.I.B., A.B., C.K., and E.P. performed research; M.R.W., P.R., O.K., P.-A.J., M.H., S.J., and A.V.-S. contributed new reagents/analytic tools; M.W., A.L., D.Z., P.W.T.C.J., H.I.B., N.S., M. Vermeulen, E.P., and R.M. analyzed data; and M.W., A.L., E.P., and R.M. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Data deposition: The data reported in this paper have been deposited in the Gene Expression Omnibus (GEO) database, https://www.ncbi.nlm.nih.gov/geo (accession no. GSE118186), and in the ProteomeXchange Consortium database via the PRIDE partner repository, https://www.ebi.ac.uk/pride/archive/ (dataset identifier PXD012547).

[1]M.W. and A.L. contributed equally to this work.

[2]To whom correspondence may be addressed. Email: eric.pasmant@parisdescartes.fr or raphael.margueron@curie.fr.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1814634116/-/DCSupplemental.

Published online March 13, 2019.

inactivation of *NF1* and *CDKN2A* tumor suppressor genes are considered to be the most significant diagnostic markers of MPNST in the revised World Health Organization 2016 classification of tumors of the central nervous system (11). Surprisingly, among 121 MPNST samples analyzed across five studies (8–10, 12, 13), no mutations were detected in *EZH1* or *EZH2* despite the high prevalence of lesions in *EED* and *SUZ12*. This mutation signature is in sharp contrast to the spectrum of PRC2 mutations found in myeloid malignancies and T-ALL, where mutations in *EZH2* occur at high frequency (Fig. 1*A*). The absence of mutations in *EZH1* and *EZH2* in MPNST raises the possibility that the enzymatic subunits might have PRC2-independent functions that are required for MPNST development. Several studies have indeed suggested that EZH1 and/or EZH2 can function independently of their enzymatic activity within PRC2 (14–16). An alternative reason for the absence of EZH2 and EZH1 mutations in MPNSTs might come from their potential functional redundancy. The two enzymes have indeed been shown to be partially redundant in several cell types (3, 17, 18).

In principle, redundancy between EZH1 and EZH2 or a PRC2-independent function for these proteins could both explain why they are not found mutated in MPNST (Fig. 1*A*). Through biochemical and genetic approaches, we unambiguously demonstrate that EZH1/2 do not regulate transcription independently of PRC2. Furthermore, we find that EZH1 and EZH2 display a remarkable degree of redundancy in the cells from which MPNST develop, providing an explanation for why alterations of the cognate genes are not selected for during MPNST development. We further provide evidence that the rate of cell proliferation is a major modulator of EZH2/EZH1 ratio and consequently the ability of EZH1 to compensate for loss of EZH2. These results suggest that context-dependent redundancies within chromatin-modifying complexes can shape mutational signatures in cancer.

Results

In the Absence of SUZ12, EZH2 Forms a Residual Complex with EED. To evaluate putative functions of EZH2 independent of an intact PRC2 core complex, we first analyzed the biochemical properties of the enzyme in the context of loss of SUZ12. For this purpose, we compared a *SUZ12*-mutated MPNST cell line (88-14) and a PRC2 wild-type cell line (STS26T). EZH2 protein accumulation appears much lower in the absence of SUZ12 (Fig. 1*B*, top Western blot; compare signal in the input fraction), consistent with previous reports showing that PRC2 constituents stabilize each other (5). To investigate the functionality of the residual EZH2 protein, we subjected nuclear extracts from 88-14 and STS26T to anion exchange chromatography followed by size exclusion chromatography (Fig. 1*B*). During anion-exchange chromatography, EZH2 elutes at 500 mM salt regardless of SUZ12 presence (Fig. 1*B*, *Top*). Following size exclusion chromatography, EZH2 and EED coeluted in both cases; however, the elution pattern is shifted toward a smaller molecular weight in the absence of SUZ12 (Fig. 1*B*, *Bottom*). These results suggest that, upon loss of SUZ12, EZH2 remains part of a smaller complex containing EED.

To determine whether EZH2 alone, or the residual complex with EED, could form new interactions in the absence of SUZ12, we characterized the interactome of EZH2 by mass spectrometry in 88-14 and STS26T cells. We overexpressed a Flag-tagged version of EZH2 in both cell lines and performed anti-Flag immunoprecipitation followed by quantitative proteomics (PXD012547, ref. 19). For the analysis of mass spectrometry data, we chose a low enrichment cutoff to avoid missing weak interactors. As expected, in PRC2 wild-type cells, EZH2 pulled down the core PRC2 components (EED, SUZ12, RBBP4, and RBBP7) along with its well-established cofactors (JARID2, AEBP2, PHF1/19, PALI1, and EPOP) (5, 20) (Fig. 1*C*, *Left*). In contrast, in *SUZ12*-mutant cells, EZH2 remained bound only to EED and lost its interaction with all accessory PRC2 subunits (Fig. 1*C*, *Right*). ZBTB17A, a potential EZH2 interactor that appeared near the enrichment cutoff, was not

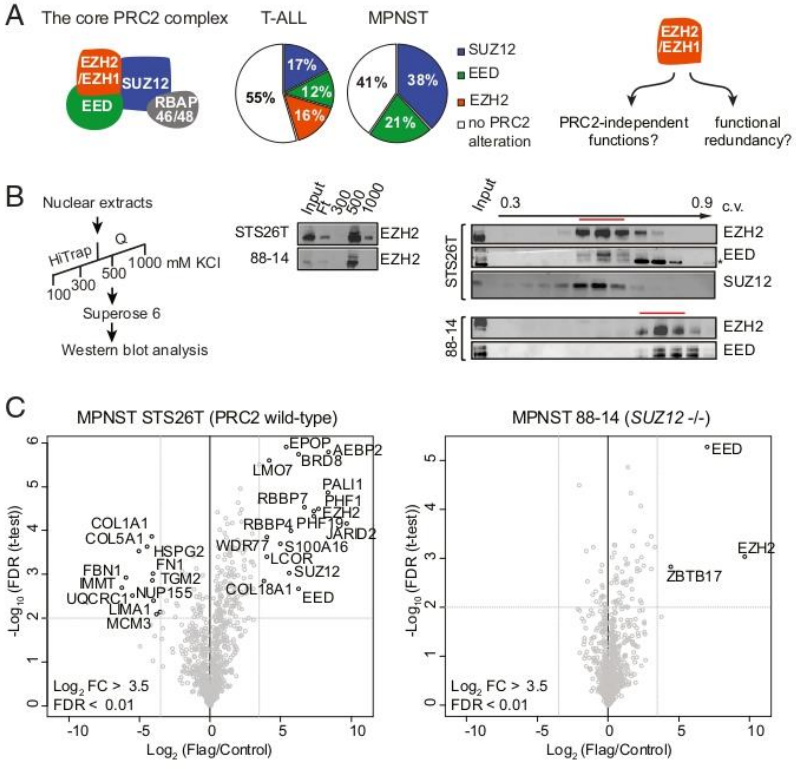


Fig. 1. Loss of SUZ12 leads to a dramatic reduction of EZH2 interactome. (*A*, *Left*) Schematic representation of the core PRC2 complex formed around EZH1 or EZH2 enzymatic subunits. (*A*, *Middle*) Pie charts displaying the relative proportions of PRC2 alterations among early T cell precursor acute lymphoblastic leukemia (T-ALL) (as reported in ref. 35) and 121 MPNST samples analyzed across five studies for which whole-genome sequencing or targeted sequencing of *EED*, *SUZ12*, *EZH1*, and *EZH2* is available. (*A*, *Right*) Alternative hypotheses explaining the absence of *EZH1/2* mutations in MPNST. (*B*, *Left*) *SUZ12*-mutated 88-14 and *SUZ12*-WT STS26T nuclear extracts were purified by successive steps summarized in the scheme. (*B*, *Middle*) Western blot analysis of EZH2 distribution in the different fractions obtained by anion exchange chromatography performed on 88-14 and STS26T nuclear extracts. (*B*, *Right*) EZH2 and EED Western blot analysis of the fractions eluted from gel filtration chromatography performed on the STS26T and 88-14 500 mM fractions. Western blot analysis of SUZ12 was additionally performed on STS26T extracts. A red line highlights the peak of elution, and the asterisk indicates an unspecific cross-reactivity of EED antibody. c.v., column volume; Ft, flow-through. (*C*) Mass spectrometry analysis of EZH2 interactome in STS26T (*Left*) and 88-14 (*Right*) stably expressing a Flag-tagged version of EZH2. Volcano plots represent mass spectrometry analysis of Flag-EZH2 cells compared with control cells.

validated by coimmunoprecipitation experiments (*SI Appendix*, Fig. S1*A*). To ascertain that the collapse of EZH2 interactome in 88-14 cells is due to the absence of SUZ12, we reexpressed SUZ12 in 88-14 cells. As shown in *SI Appendix*, Fig. S1*B*, the binding of the PRC2 cofactor AEBP2 to EZH2 was restored in the presence of SUZ12. In addition, expression of other key PRC2 cofactors was equivalent in 88-14 and STS26T cells (*SI Appendix*, Fig. S1*C*), indicating that these two cell lines have a similar environment for the regulation of PRC2. The massive loss of interaction between EZH2 and PRC2 cofactors in *SUZ12*-mutant cells is in agreement with recent in vitro evidence suggesting that SUZ12 acts as a platform for the recruitment of these subunits (21, 22).

Altogether, these analyses show that loss of SUZ12 destabilizes the PRC2 interactome, leaving EZH2 and EED as the only members of a residual PRC2 complex.

EZH1/2 Do Not Regulate Transcription Independently of PRC2. Several studies have suggested that EZH2 might function independently of canonical PRC2 activity (14–16). *SUZ12*-mutant MPNST cell lines represent an ideal system to rigorously investigate such noncanonical functions because the canonical PRC2 function is inactivated. Since EZH1, the paralogous enzyme of EZH2, can in some instances compensate for loss of EZH2, we assessed the functionality of both enzymes together. We analyzed the impact of inhibiting EZH1/2 enzymatic activity on cell proliferation, using the small-molecule inhibitor UNC1999 or its inactive analog UNC2400 (23). As shown in Fig. 2*A*, treatment of two different *SUZ12*-null MPNST cell lines did not impact cell proliferation. However, this observation does not exclude a role for the residual PRC2 complex in controlling gene expression and/or functions of EZH2 that do not rely on its catalytic activity.

To address this question, we genetically inactivated *EZH1* and *EZH2* in the *SUZ12*-null 88-14 cell line. Loss of EZH2 in three independent clones was confirmed by Western blot (Fig. 2*B*). We introduced frameshift-inducing mutations in the *EZH1* gene to exclude a potential compensation upon loss of EZH2 (*SI Appendix*, Fig. S2*A*). These mutations do not affect proliferation as shown by the lack of consequences of restoring EZH2 expression in 88-14 *EZH1/2* double knockout (dKO) (*SI Appendix*, Fig. S2*B*). We performed RNA sequencing on *EZH1/2* wild-type and *EZH1/2* dKO clones. Strikingly, with the exception of *EZH2*, no genes were found significantly differentially expressed [false discovery rate (FDR) < 0.05] between control and *EZH1/2* dKO cells (GSE118186, ref. 24) (Fig. 2*C*).

These results strongly argue against a PRC2-independent function for EZH2 in the context of MPNST.

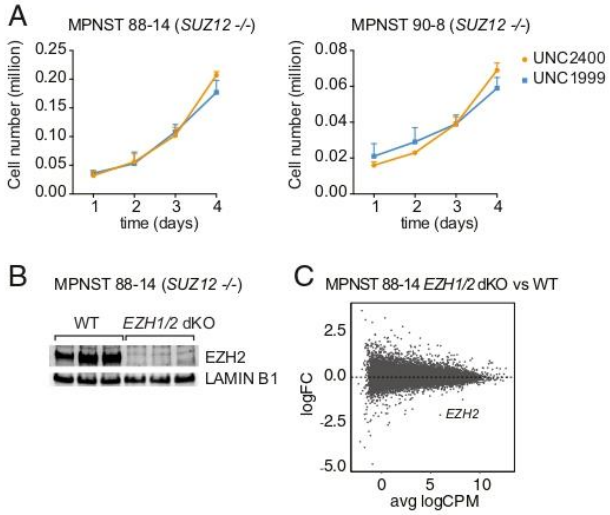


Fig. 2. EZH2 does not retain a specific function in the absence of SUZ12 in MPNST cells. (*A*) Proliferation assays in two *SUZ12*-null MPNST cell lines in presence of UNC1999 dual EZH1/2 inhibitor or of UNC2400, its control inactive compound (mean ± SD; $n = 3$). (*B*) Western blot analysis EZH2 protein expression in three independent wild-type and *EZH1/2* dKO 88-14 clones. (*C*) Scatterplot showing $\log_2$fold-change (logFC) expression between wild-type and *EZH1/2* dKO 88-14 cells versus average $\log_2$ counts per million (logCPM). *EZH2*, the only differentially expressed transcript, is highlighted in red.

EZH2 Functions as Part of a Canonical PRC2 Complex in Androgen-Independent LNCaP-abl Cells. Our findings in the context of MPNST contrast with those reported in androgen-independent prostate cancer (AIPC), where biochemical and functional evidence points to a PRC2-independent role for EZH2, mediating gene activation as part of a distinct complex comprising the androgen receptor (16).

To understand the basis of this discrepancy, we repeated the biochemical characterization of PRC2 in the LNCaP-abl cell line, an androgen-independent derivative of the LNCaP prostate cancer cell line used as model of AIPC (16, 25). In contrast to Xu et al., we observed coelution of EZH2 with EED following size exclusion chromatography, a pattern identical to that observed in STS26T cells (Fig. 3*A* and *SI Appendix*, Fig. S3*A*; compare with Fig. 1*B*). Analysis of EZH2 migration on a native gel further confirmed that it is found in a single high–molecular-weight complex slightly smaller than the recombinant complex where all subunits are tagged (*SI Appendix*, Fig. S3*B*). Moreover, mass spectrometry analysis of the EZH2 interactome in LNCaP-abl cells recovered all known core and accessory PRC2 subunits but did not reveal additional partners such as the androgen receptor despite the low stringency used in the analysis (PXD012547, ref. 19) (Fig. 3*B*). These biochemical analyses therefore do not support the existence of measurable noncanonical composition of PRC2 in this AIPC cell line. Of note, we obtained similar findings for the OVCAR8 cell line (*SI Appendix*, Fig. S3*C*), in which EZH2 has also been reported to have an unusual elution pattern on a size exclusion column (26).

These results prompted us to further investigate EZH2 contribution to LNCaP-abl growth. We measured cell proliferation upon EZH2 enzymatic inhibition since the PRC2-independent function of EZH2 has been shown to require an intact catalytic domain (16). We pretreated cells with UNC1999 or UNC2400 and performed cell growth assays beginning after either 4 or 15 d of continued treatment. Efficient inhibition of EZH1/2 catalytic activity was verified by Western blot for the trimethylated form of H3K27 (H3K27me3; *SI Appendix*, Fig. S3*D*). In contrast to the dramatic effects observed with siRNA against EZH2 on cell growth as early as 48 h after transfection, pharmacological inhibition of PRC2 only impaired LNCaP-abl growth after the first week of treatment (Fig. 3*C*). This delay has been reported in other models (23, 27, 28) and attributed to the stability of H3K27me3 and the time required to achieve full dilution of the histone mark through cell divisions.

Importantly, the transcriptomic data generated following siEZH2 treatment (16) revealed a prominent cell-proliferation signature among down-regulated genes (i.e., genes that are proposed to be regulated by EZH2 in a noncanonical fashion; *SI Appendix*, Fig. S3*E*). Indeed, following PRC2 inhibition, expression of cell proliferation markers is diminished but only at 10 d posttreatment (Fig. 3*D*, *Top*). This is the same kinetics as the one observed for the derepression of classical PRC2 target genes (Fig. 3*D*, *Bottom*) raising the possibility that down-regulation of cell proliferation genes is an indirect consequence of the more global transcriptional alterations resulting from EZH2 inhibition.

Thus, together with our biochemical analyses, the kinetics of cell growth suppression and of gene expression changes upon PRC2 inhibition suggests a predominantly canonical function for EZH2 in LNCaP-abl cells.

Pronounced Redundancy Between EZH1 and EZH2 in Neurofibroma Cells. Considering the lack of evidence for PRC2-independent role of EZH2 that would explain why it is not found to be mutated in MPNST, we investigated the alternative hypothesis that the

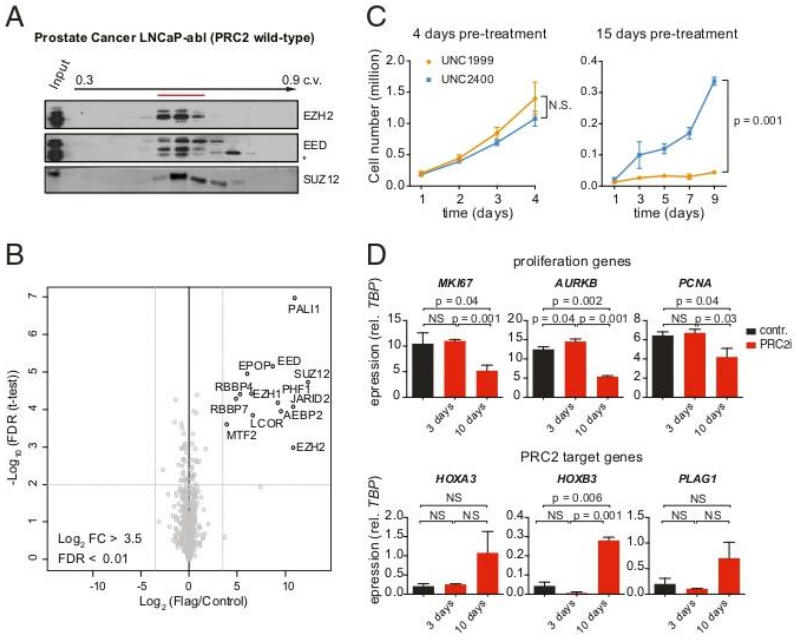


Fig. 3. EZH2 acts as part of a canonical PRC2 complex in LNCaP-abl cells. (*A*) Western blot analysis of EZH2 and EED in the fractions eluted from gel filtration chromatography performed on the 500 mM fraction. A red line highlights the peak of elution, and the asterisk indicates an unspecific cross-reactivity of EED antibody. c.v., column volume. (*B*) Mass spectrometry analysis of EZH2 interactome in LNCaP-abl cells stably expressing Flag-tagged version of EZH2. (*C*) Proliferation assay in LNCaP-abl treated with 1 μM UNC1999 or UNC2400 from the indicated times after the beginning of drug treatment. Mean ± SD; $n = 2$; *P* values from two-tailed *t* tests on the final time points are shown. (*D*) RT-qPCR analysis of proliferation markers (*Top*) and classical PRC2 target genes (*Bottom*) in control-treated (UNC2400) cells or at different times after PRC2 inhibitor treatment. Mean ± SD; $n = 2$; *P* values from one-tailed *t* tests are shown for differences that are statistically significant. NS, not significant.

absence of mutations affecting EZH2 might reflect compensation by EZH1. We inactivated *EZH1* and *EZH2* separately or in combination in the ipNF05.5 cell line derived from a plexiform neurofibroma (29), corresponding to the tumor type from which MPNSTs arise. The resulting mutant cells were compared with *SUZ12*-mutant cells derived from the same cell line (GSE118186, ref. 24) (Fig. 4*A*). Loss of SUZ12 in ipNF05.5 cells led to transcriptional up-regulation of 813 genes (*SI Appendix*, Fig. S4*A*; FDR < 0.05). As expected, a majority of these genes are marked by H3K27me3 in wild-type ipNF05.5 cells (*SI Appendix*, Fig. S4*B*). Gene ontology analysis indicated that up-regulated genes were significantly enriched for developmental regulators (*SI Appendix*, Fig. S4*C*), as has been found in PRC2-mutant MPNST samples (9).

We then analyzed the impact of individual or combined loss of EZH1 and EZH2 on the methylation of H3K27 (Fig. 4*B*; also see *SI Appendix*, Fig. S4*D* for quantification of Western blot signals). While loss of EZH1 did not impact global levels of H3K27me1, me2, or me3, loss of EZH2 led to a marked reduction of H3K27me3 with only modest effects on H3K27me1 and H3K27me2. Combined loss of both enzymes or loss of SUZ12 leads to a comparable acute loss of all three methylation levels. We next assessed transcriptional changes by RNA-seq (Fig. 4*C* and *SI Appendix*, Fig. S4 *E* and *F*). Strikingly, deletions of either *EZH1* or *EZH2* caused only subtle changes in gene expression, with no differentially expressed genes in *EZH1*-mutant cells and only 11 in *EZH2*-mutant cells. In contrast, combined loss of EZH1 and EZH2 led to up-regulation of 629 genes, indicative of a high degree of redundancy between the two enzymes. Consistently, inhibition of PRC2 activity in *EZH2* KO ipNF05.05 using the general PRC2 inhibitor A-395 (30) led to a robust de-repression of PRC2 target genes (*SI Appendix*, Fig. S4*H*). These data support the hypothesis that the absence of *EZH2* (or *EZH1*) mutations in MPNST is a consequence of the high redundancy between the two enzymes in neurofibroma cells.

Remarkably, transcriptional changes in *EZH1*/*EZH2* dKO and *SUZ12* KO were highly correlated (GSE118186, ref. 24) (Fig. 4*C* and *SI Appendix*, Fig. S4*I*). Direct comparison of the two mutant conditions did not uncover any significantly differentially expressed genes (*SI Appendix*, Fig. S4*J*), demonstrating that combined loss of EZH1 and EZH2 is equivalent to loss of SUZ12. Together with our observations in the MPNST and AIPC models, these results in neurofibroma cells argue against a PRC2-independent function for EZH1/2.

Redundancy Between EZH1 and EZH2 Is Modulated by Cell Proliferation Rate. Previous studies have found that functional compensation between EZH1 and EZH2 is context dependent, depending on tissue type and developmental stages (17, 31). Furthermore, as mentioned above, mutations in *EZH2* are selected for in T-ALL and in myeloid malignancies, suggesting that EZH1 cannot fully compensate for loss of EZH2 in these cell types. The circumstances under which EZH1 and EZH2 compensate for each other remain unclear. We and others have previously shown that EZH2 expression is driven by cell proliferation (3, 18, 32), a process that ensures H3K27me3 homeostasis (18). As shown in *SI Appendix*, Fig. S5*A*, analysis of publicly available RNA-seq data from The Cancer Genome Atlas database indicates that the positive correlation between *EZH2* transcript levels and cell proliferation (assessed by *MKI67* proliferation marker) is a general property that extends across various cancer types. In contrast, *EZH1* levels show no positive correlation to cell proliferation, suggesting that the EZH2/EZH1 ratio and hence redundancy between the two enzymes is mainly controlled by cell proliferation rate.

To directly evaluate the link between EZH2/EZH1 ratio and cell proliferation in tumor types subject to mutation affecting PRC2 genes, we analyzed tumor samples of autonomous thyroid adenoma (ATA) and plexiform neurofibroma (PNF), two tumor types in which *EZH2* is never found mutated, as well as myelodysplastic syndrome (MDS) and follicular lymphoma (FL), in which *EZH2* mutations occur. Using digital droplet PCR, we quantitatively measured *EZH1*, *EZH2*, and *MKI67* transcript abundance. Fig. 5*A* shows the strong correlation between *EZH2*/*EZH1* ratio and *MKI67* expression (Spearman $r = 0.7$, $P < 0.0001$). Samples are distributed in two separate clusters, ATAs and PNFs forming a cluster characterized by low *MKI67* expression and low *EZH2*/*EZH1* ratio, while MDSs and FLs samples form a cluster characterized by high *MKI67* and high *EZH2*/*EZH1* ratio. Thus, these observations reveal a striking association between proliferation status, *EZH2*/*EZH1* ratio, and the occurrence of *EZH2* mutations, the latter being found only in highly proliferative tumors.

We next sought to directly assess whether redundancy between EZH1 and EZH2 is indeed alleviated upon higher proliferation. To

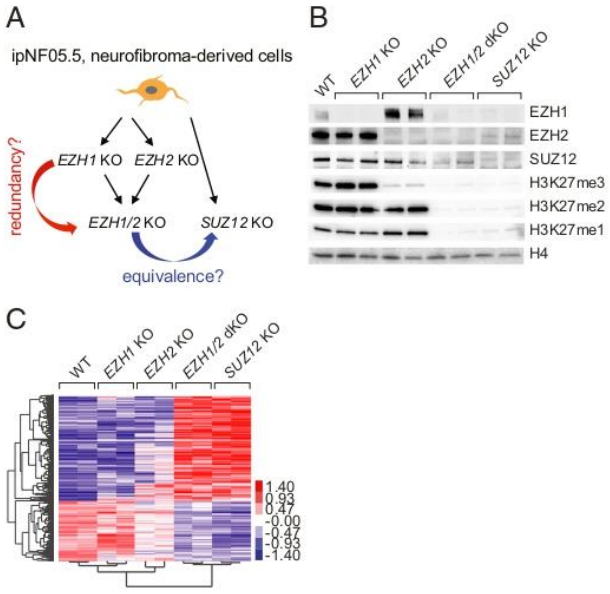


Fig. 4. Acute redundancy between EZH1 and EZH2 in a model of plexiform neurofibroma. (*A*) Schematic of the different isogenic KO cell lines established from the ipNF05.5 plexiform neurofibroma cell line and of the experimental rationale. (*B*) Western blot showing EZH1, EZH2, SUZ12, and the various degrees of H3K27 methylation in wild-type ipNF05.5 and the different mutant clones indicated on *Top*. Each KO condition is represented by two independent clones. Histone H4 is used as a loading control. (*C*) Heatmap of centered, log-transformed values (counts per million) of all differentially expressed genes, each KO condition versus WT. Genes and samples were clustered using Spearman rank correlation.

this aim, we compared the consequence of deleting *EZH2* in ipNF05.5 cells and HAP1 cells, which have a much higher proliferation rate (12-h doubling time compared with 30 h for ipNF05.5 cells). The level of EZH1 is similar between the two cell lines, but EZH2 is much more abundant in HAP1 cells than in ipNF05.5 cells and is paralleled by a high level of the PCNA proliferation marker (Fig. 5*B*). Interestingly, substantial levels of H3K27me2 and me3 remain in *EZH2* KO ipNF05.5 cells, while loss of EZH2 in HAP1 cells results in acute loss of H3K27me2 and me3, similar to that obtained upon deletion of *EED* (Fig. 5*B*; also see *SI Appendix*, Fig. S5*B* for quantification of Western blot signals). Accordingly, loss of EZH2 in HAP1 cells led to transcriptional derepression of a number of genes, similar to KO of the core PRC2 component EED, but has moderate transcriptional consequences in ipNF05.5 cells (Fig. 5*C*). The differential effect of deleting *EZH2* in ipNF05.5 and HAP1 cells is further illustrated at genes that are regulated by PRC2 in both cell lines (*SI Appendix*, Fig. S5*C*).

To further ascertain that the association between cell proliferation and EZH2/EZH1 redundancy is not a bias resulting from analyzing different cell types, we investigated the impact of direct manipulation of cell proliferation rate in the context of a defined cell type. We performed a KO of *EZH2* in an immortalized Schwann cell line (29) and grew wild-type and *EZH2* KO cells in either low-serum (slow-proliferation) or high-serum (high-proliferation) medium. As expected, upon high proliferation, EZH2 levels strongly increased while EZH1 expression remain constant (Fig. 5*D*). Interestingly, in the absence of EZH2, higher proliferation led to a decrease of H3K27me2/me3 levels (Fig. 5*D*; also see *SI Appendix*, Fig. S5*D* for quantification of Western blot signals). This experiment thus demonstrates that EZH1's ability to compensate for loss of EZH2 is inversely proportional to the rate of cell proliferation.

Altogether, our analyses show that EZH1 cannot compensate for loss of EZH2 under high proliferation, suggesting that proliferation is a major factor underlying the redundancy between EZH1 and EZH2. We propose that the proliferative index is a key constraint underlying the PRC2 mutation pattern that is selected for in the course of tumorigenesis (*SI Appendix*, Fig. S5*E*). In tumors characterized by a low proliferation index such as PNFs, mutations in EZH2 will not be selected for because of the redundancy with EZH1. In tumor types characterized by a much higher proliferative index such as myeloid malignancies and T-ALL, EZH2 becomes predominant relative to EZH1, and thus loss of EZH2 can be selected for. Interestingly, this simple model also accounts for gain-of-function mutations that selectively occur on EZH1 in the slow proliferating ATAs or on EZH2 in the more proliferative FL and diffuse large B cell lymphoma types of lymphomas.

Discussion

The PRC2 complex is diverted from its normal function in cancer through defined tumor-type specific mutations. These alterations have been suggested to entail canonical as well as noncanonical functions of EZH2, the main PRC2 enzymatic subunit. However, we currently lack a precise understanding of how mutations found in cancer affect each of these activities, thus limiting our ability to develop rational therapeutic approaches. In this study, we investigated the relative contribution of canonical versus noncanonical activities of EZH2 in the regulation of gene expression. MPNST cells represent an ideal system to study such a noncanonical activity since the canonical PRC2 function is absent. In such a context, we provide compelling biochemical and genetic evidence that EZH2 does not regulate transcription. PRC2-independent functions could still be at play in certain cell types; however, our biochemical analysis of EZH2 and the kinetics of cell growth upon EZH2 inhibition in the LNCaP-abl model cell line are more consistent with a canonical function of the enzyme.

We uncover a pronounced functional redundancy between EZH1 and EZH2 in MPNST precursor cells. Combined loss of

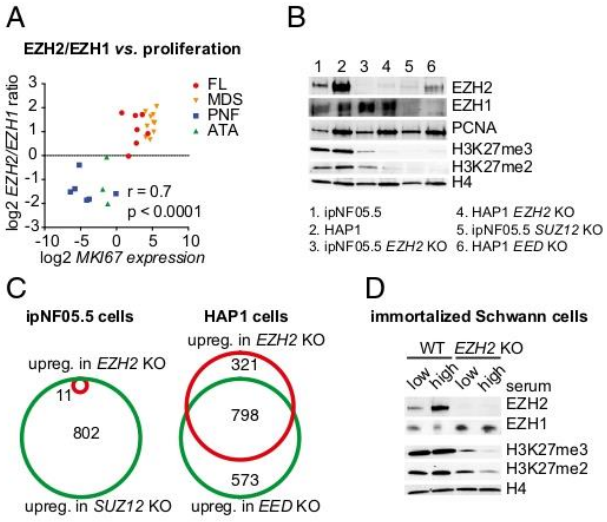


Fig. 5. Redundancy between EZH1 and EZH2 is dependent on the proliferation rate. (*A*) Scatterplot of *EZH2/EZH1* transcript ratio versus *MKI67* transcript abundance as determined by digital droplet PCR in the tumor samples indicated on the *Right*. (*B*) Western blot showing the comparative abundance of EZH1, EZH2, the PCNA proliferation marker, and H3K27me2 and H3K27me3 in ipNF05.5 versus HAP1 cells in the indicated wild-type or KO conditions. H4 is used as a loading control. (*C*) Venn diagrams showing the overlap between genes up-regulated upon complete PRC2 inactivation (*SUZ12* or *EED* KO) and upon *EZH2* KO in ipNF05.5 and the highly proliferative HAP1 cells. (*D*) Western blot showing EZH1, EZH2, H3K27me2, and H3K27me3 levels in wild-type or *EZH2* KO immortalized Schwann cells grown in low proliferation medium (low serum) or high proliferation medium (high serum).

both enzymes leads to gene expression changes that are indistinguishable from those induced by loss of SUZ12. These findings strongly suggest that the signature of PRC2 mutations in MPNST, that is, the absence of mutations in EZH2/EZH1, stems from the high redundancy between the two enzymes. In this context, the probability of inactivating PRC2 through biallelic mutation of *EZH1* and *EZH2* (four alleles) would be far smaller than mutation of *SUZ12* or *EED* (two alleles only at most; one allele of *SUZ12* often being inactivated simultaneously with the driving mutation in *NF1* by a large deletion encompassing both genes).

Our analyses also show that the rate of cell proliferation, which positively controls the expression of EZH2 but not EZH1, is a key parameter modulating the redundancy between the two enzymes, explaining why mutations in EZH2 are favored in high proliferating tumor types. Nonetheless, other independent cues are also likely to participate in modulating the balance between EZH2 and EZH1 as recently shown for EZH1 expression during male germ cell development (31). As in the case of PRC2, mutations in genes encoding subunits of other chromatin-modifying complexes such as SWI/SNF and COMPASS show a certain degree of tumor type specificity (33, 34). It will be interesting to investigate whether similar context-dependent redundancies within these chromatin-modifying complexes underlie such mutation patterns. In conclusion, our study addresses key issues regarding canonical activities of the PRC2 complex and provides insights into the characteristic spectrum of PRC2 mutations found in different cancers.

Materials and Methods

Cell Culture. *SUZ12* mutated 88-14 and *SUZ12* wild-type STS26T cell lines were kindly provided by Nancy Ratner, Cincinnati Children's Hospital Medical Center, Cincinnati. The LNCaP-abl cell line was kindly provided by Zoran Culig, Innsbruck Medical University, Innsbruck, Austria. ipNF05.5 plexiform neurofibroma cell line and immortalized Schwann cell line are described in ref. 29. HAP1 cells were kindly provided by T. Brummelkamp, Oncode Institute, Amsterdam. OVCAR8 cells were kindly provided by Fatima Mechta-Grigoriou, Institut Curie, Paris. Additional details cell culture conditions, transfection, constitutive knockouts, proliferation assays, and downstream biochemical analyses are provided in *SI Appendix, Supplementary Methods*.

RNA Extraction, RT-qPCR, Digital Droplet PCR, RNA Sequencing, and Data Analysis. Total RNA was isolated using TRIzol-chloroform extraction and isopropanol precipitation. Additional details on RT-qPCR, digital droplet PCR, RNA sequencing, and data analysis are provided in *SI Appendix, Supplementary Methods*.

Chromatin Immunoprecipitation. Chromatin immunoprecipitation (ChIP) was performed as described previously (3). Additional details on ChIP sequencing and data analysis are provided in *SI Appendix, Supplementary Methods*.

Patients and Samples. All patients provided informed consent, and the study was approved by institutional review board and local ethical committees. Project ID CPP17/79, A0296746, and 2015-08-11DC were reviewed by Cochin Hospital institutional review board and CPP Ile-de-France 2 ethics committee APHP, Paris, and project ID BS#2017-311 by the Groupe Thématique de Travail–Hematology Section, Institut Curie, institutional review board and ethics committee of the Hospital Group, Institut Curie, Paris. Additional details on patient samples are provided in *SI Appendix, Supplementary Methods*.

Data Access. Next-generation sequencing (NGS) data have been deposited in the GEO database (accession no. GSE118186). Mass spectrometry data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD012547.

ACKNOWLEDGMENTS. We thank Pascale Gilardi, Daniel Holoch, and members of the R.M. laboratory for comments on the manuscript. Work in the laboratory of R.M. is supported by the "Association pour la Recherche sur le Cancer" and the Labex "Development, Epigenetics, Epigenetics, and Lifetime Potential." R.M. and E.P. were supported by Institut Thématique Multi-Organisme Cancer (Grant EpiNF1). The M.V. lab is part of the Oncode Institute, which is partly funded by the Dutch Cancer Society (KWF). A.L. was a recipient of a fellowship from INSERM. High-throughput sequencing was performed by the NGS platform of the Institut Curie, supported by Grants ANR-10-EQPX-03 and ANR10-INBS-09-08 from the Agence Nationale de le Recherche (Investissements d'Avenir) and by the Canceropôle Ile-de-France.

1. Shah MA, Denton EL, Arrowsmith CH, Lupien M, Schapira M (2014) A global assessment of cancer genomic alterations in epigenetic mechanisms. *Epigenetics Chromatin* 7:29.
2. Workman P, Al-Lazikani B (2013) Drugging cancer genomes. *Nat Rev Drug Discov* 12: 889–890.
3. Margueron R, et al. (2008) Ezh1 and Ezh2 maintain repressive chromatin through different mechanisms. *Mol Cell* 32:503–518.
4. Shen X, et al. (2008) EZH1 mediates methylation on histone H3 lysine 27 and complements EZH2 in maintaining stem cell identity and executing pluripotency. *Mol Cell* 32:491–502.
5. Holoch D, Margueron R (2017) Mechanisms regulating PRC2 recruitment and enzymatic activity. *Trends Biochem Sci* 42:531–542.
6. Pengelly AR, Copur Ö, Jäckle H, Herzig A, Müller J (2013) A histone mutant reproduces the phenotype caused by loss of histone-modifying factor Polycomb. *Science* 339:698–699.
7. Wassef M, Margueron R (2017) The multiple facets of PRC2 alterations in cancers. *J Mol Biol* 429:1978–1993.
8. De Raedt T, et al. (2014) PRC2 loss amplifies Ras-driven transcription and confers sensitivity to BRD4-based therapies. *Nature* 514:247–251.
9. Lee W, et al. (2014) PRC2 is recurrently inactivated through EED or SUZ12 loss in malignant peripheral nerve sheath tumors. *Nat Genet* 46:1227–1232.
10. Zhang M, et al. (2014) Somatic mutations of SUZ12 in malignant peripheral nerve sheath tumors. *Nat Genet* 46:1170–1172.
11. Sahm F, Reuss DE, Giannini C (2018) WHO 2016 classification: Changes and advancements in the diagnosis of miscellaneous primary CNS tumours. *Neuropathol Appl Neurobiol* 44:163–171.
12. Brohl AS, Kahen E, Yoder SJ, Teer JK, Reed DR (2017) The genomic landscape of malignant peripheral nerve sheath tumors: Diverse drivers of Ras pathway activation. *Sci Rep* 7:14992.
13. Sohier P, et al. (2017) Confirmation of mutation landscape of NF1-associated malignant peripheral nerve sheath tumors. *Genes Chromosomes Cancer* 56:421–426.
14. Gonzalez ME, et al. (2014) EZH2 expands breast stem cells through activation of NOTCH1 signaling. *Proc Natl Acad Sci USA* 111:3098–3103.
15. Lee ST, et al. (2011) Context-specific regulation of NF-κB target gene expression by EZH2 in breast cancers. *Mol Cell* 43:798–810.
16. Xu K, et al. (2012) EZH2 oncogenic activity in castration-resistant prostate cancer cells is Polycomb-independent. *Science* 338:1465–1469.
17. Ezhkova E, et al. (2011) EZH1 and EZH2 cogovern histone H3K27 trimethylation and are essential for hair follicle homeostasis and wound repair. *Genes Dev* 25:485–498.
18. Wassef M, et al. (2015) Impaired PRC2 activity promotes transcriptional instability and favors breast tumorigenesis. *Genes Dev* 29:2547–2562.
19. Jansen PTWC, Margueron R (2019) Data from "EZH1/2 function mostly within canonical PRC2 and exhibit proliferation-dependent redundancy that shapes mutational signatures in cancer." PRIDE. Available at www.ebi.ac.uk/pride/archive/projects/PXD012547. Deposited January 29, 2019.
20. Conway E, et al. (2018) A family of vertebrate-specific Polycombs encoded by the LCOR/LCORL genes balance PRC2 subtype activities. *Mol Cell* 70:408–421.e8.
21. Chen S, Jiao L, Shubbar M, Yang X, Liu X (2018) Unique structural platforms of Suz12 dictate distinct classes of PRC2 for chromatin binding. *Mol Cell* 69:840–852.e5.
22. Kasinath V, et al. (2018) Structures of human PRC2 with its cofactors AEBP2 and JARID2. *Science* 359:940–944.
23. Konze KD, et al. (2013) An orally bioavailable chemical probe of the lysine methyltransferases EZH2 and EZH1. *ACS Chem Biol* 8:1324–1334.
24. Wassef M, et al. (2019) Data from "EZH1/2 function mostly within canonical PRC2 and exhibit proliferation-dependent redundancy that shapes mutational signatures in cancer." GEO. Available at https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE118186. Deposited August 6, 2018.
25. Culig Z, et al. (1999) Switch from antagonist to agonist of the androgen receptor bicalutamide is associated with prostate tumour progression in a new model system. *Br J Cancer* 81:242–251.
26. Wan L, et al. (2018) Phosphorylation of EZH2 by AMPK suppresses PRC2 methyltransferase activity and oncogenic function. *Mol Cell* 69:279–291.e5.
27. Bradley WD, et al. (2014) EZH2 inhibitor efficacy in non-Hodgkin's lymphoma does not require suppression of H3K27 monomethylation. *Chem Biol* 21:1463–1475.
28. Knutson SK, et al. (2012) A selective inhibitor of EZH2 blocks H3K27 methylation and kills mutant lymphoma cells. *Nat Chem Biol* 8:890–896.
29. Li H, Chang L-J, Neubauer DR, Muir DF, Wallace MR (2016) Immortalization of human normal and NF1 neurofibroma Schwann cells. *Lab Invest* 96:1105–1115.
30. He Y, et al. (2017) The EED protein-protein interaction inhibitor A-395 inactivates the PRC2 complex. *Nat Chem Biol* 13:389–395.
31. Mu W, Starmer J, Shibata Y, Yee D, Magnuson T (2017) EZH1 in germ cells safeguards the function of PRC2 during spermatogenesis. *Dev Biol* 424:198–207.
32. Bracken AP, et al. (2003) EZH2 is downstream of the pRB-E2F pathway, essential for proliferation and amplified in cancer. *EMBO J* 22:5323–5335.
33. Masliah-Planchon J, Bièche I, Guinebretière J-M, Bourdeaut F, Delattre O (2015) SWI/SNF chromatin remodeling and human malignancies. *Annu Rev Pathol* 10:145–171.
34. Meeks JJ, Shilatifard A (2017) Multiple roles for the MLL/COMPASS family in the epigenetic regulation of gene expression and in cancer. *Annu Rev Cancer Biol* 1:425–446.
35. Zhang J, et al. (2012) The genetic basis of early T-cell precursor acute lymphoblastic leukaemia. *Nature* 481:157–163.

The American Journal of Pathology, Vol. 188, No. 10, October 2018

ELSEVIER

The American Journal of PATHOLOGY

ajp.amjpathol.org

TUMORIGENESIS AND NEOPLASTIC PROGRESSION

Medullary Breast Carcinoma, a Triple-Negative Breast Cancer Associated with *BCLG* Overexpression

Pierre Romero,[*] Vanessa Benhamo,[†‡] Gabrielle Deniziaut,[*] Laetitia Fuhrmann,[*†] Frédérique Berger,[§] Elodie Manié,[†] Jaydutt Bhalshankar,[†] Sophie Vacher,[¶] Cécile Laurent,[‡] Elisabetta Marangoni,[‡] Nadège Gruel,[†‡] Gaëtan MacGrogan,[‖] Roman Rouzier,[**] Olivier Delattre,[†] Tatiana Popova,[†] Fabien Reyal,[‡**] Marc-Henri Stern,[*†] Dominique Stoppa-Lyonnet,[*†,††] Caterina Marchiò,[*‡‡] Ivan Bièche,[¶§§] and Anne Vincent-Salomon[*†]

From the Departments of Pathology, Translational Research,[‡] and Surgery,** the Pharmacogenomics Unit,[¶] Department of Genetics, INSERM U934,[†] and the Unit of Biometry,[§] INSERM U900, Institut Curie, PSL Research University, Paris, France; the Department of Biopathology,[‖] Institut Bergonié, Bordeaux, France; Sorbonne Paris Cité[††] and EA 7331,[§§] University Paris Descartes, Paris, France; and the Institute of Pathology at the Department of Medical Sciences,[‡‡] University of Turin, Turin, Italy

Accepted for publication June 19, 2018.

Address correspondence to Pierre Romero, M.D., Department of Pathology, Institut Curie, 26 rue d'Ulm, 75248 Paris Cedex 05, France. E-mail: pierre.romero@curie.fr.

Medullary breast carcinoma (MBC) is a rare subtype of triple-negative breast cancer with specific genomic features within the spectrum of basal-like carcinoma (BLC). In this study of 19 MBCs and 36 non-MBC BLCs, we refined the transcriptomic and genomic knowledge about this entity. Unsupervised and supervised analysis of transcriptomic profiles confirmed that MBC clearly differs from non-MBC BLC, with 92 genes overexpressed and 154 genes underexpressed in MBC compared with non-MBC BLC. Immunity-related pathways are the most differentially represented pathways in MBC compared with non-MBC BLC. The proapoptotic gene *BCLG* (official name *BCL2L14*) is by far the most intensely overexpressed gene in MBC. A quantitative RT-PCR validation study conducted in 526 breast tumors corresponding to all molecular subtypes documented the specificity of *BCLG* overexpression in MBC, which was confirmed at the protein level by immunohistochemistry. We also found that most MBCs belong to the immunomodulatory triple-negative breast cancer subtype. Using pan-genomic analysis, it was found that MBC harbors more losses of heterozygosity than non-MBC BLC. These observations corroborate the notion that MBC remains a distinct entity that could benefit from specific treatment strategies (such as deescalation or targeted therapy) adapted to this rare tumor type. *(Am J Pathol 2018, 188: 2378–2391; https://doi.org/10.1016/j.ajpath.2018.06.021)*

Medullary breast carcinoma (MBC) is a rare morphologic subtype of invasive breast cancer, accounting for approximately 2% of invasive breast cancer.[1] Most MBC belongs to the immunohistochemical (IHC) subgroup of triple-negative breast cancer (TNBC), and at the transcriptomic level, MBC usually displays basal-like genetic features.[2] In previous studies, we found that, although it shares common genomic alterations with non-MBC basal-like carcinoma (BLC), MBC is a distinct genomic entity within the BLC spectrum, harboring a higher rate of *TP53* mutations.[3,4] Moreover, the prognosis of MBC appears to be slightly better than that of grade-matched invasive carcinoma of no special type, despite its aggressive histopathologic features.[5] To define more clearly the molecular events that characterize MBC as a morphologic and genomic subgroup of BLC, we performed a comprehensive retrospective study comparing MBC with non-MBC BLC through a transcriptomic analysis and a more detailed genomic analysis than in our previous work.[3]

Supported by the Grand Prix Ruban Rose 2012 (A.V.-S.) and in part by a grant from the Mayent-Rothschild Foundation (C.M.).

Disclosures: A patent based on the data presented in this report is pending (US Patent application US20170260588A1; T.P. and M.-H.S. are the named inventors). The patent is currently licensed to Myriad Genetics.

https://doi.org/10.1016/j.ajpath.2018.06.021

Materials and Methods

Patients, Tumors, Cell Lines, and Xenografts

Tumors were retrospectively selected from two groups of patients on the basis of the availability of frozen samples and formalin-fixed, paraffin-embedded samples for determination of transcriptomic and genomic profiles and IHC, respectively. Samples with <50% of tumor cells were excluded from the study. The first group consisted of 19 MBCs, and the second group consisted of 36 grade III non-MBC BLCs (retrieved from Institut Curie Paris and Institut Bergonié in France). Experiments were performed in accordance with the French Bioethics Law 2004-800 and the National Institute of Cancer Ethics Charter. Morphologic and IHC assessments of diagnostic criteria in MBC were performed as previously described, in accordance with the criteria defined by Ridolfi et al.[3,6] Non-MBC BLC was defined using morphology and IHC as follows: absence of MBC morphologic criteria and lack of expression of estrogen receptor (ER), progesterone receptor (PR), and HER2 (ie, triple-negative phenotype), with the expression of at least one of the following markers: keratin 5/6, keratin 14, epidermal growth factor receptor, and KIT. Tumor-infiltrating lymphocytes (TILs) assessment in non-MBC BLC was performed by two independent readers (P.R., C.M.) as recommended by Salgado et al.[7] Clinicopathological data are detailed in Supplemental Table S1.

Transcriptomic data validation analysis was performed with RNA tumor samples extracted from a series of 526 primary unilateral invasive breast tumors (retrieved from Institut Curie, Paris, France) between 1978 and 2008. Complete clinical, histologic, and laboratory data were available. Patients were managed by primary tumor excision without neoadjuvant chemotherapy or radiotherapy. ER, PR, and HER2 protein status was determined by biochemical methods (dextran-coated charcoal method, enzyme immunoassay, or IHC) and was confirmed by real-time quantitative RT-PCR (RT-qPCR) assays. The population was divided into four groups that reflected the intrinsic molecular classification: two luminal subtypes, luminal A [HR^+ (ER^+ or PR^+)/$HER2^-$ ($n = 295$)] and luminal B [HR^+ (ER^+ or PR^+)/$HER2^+$ ($n = 58$)]; $HER2^+$ subtype [HR^- (ER^- and PR^-)/$HER2^+$ ($n = 72$)]; and triple-negative (TN) subtype [HR^- (ER^- and PR^-)/$HER2^-$ ($n = 101$)]. Sixteen normal breast tissue specimens from women undergoing cosmetic breast surgery were used as sources of normal RNA.

For validation of the gene of interest, 38 RNA samples extracted from cell lines obtained from the ATCC (Manassas, VA) were tested: seven nontumor cell lines, 17 TN tumor cell lines, 10 ER^+ tumor cell lines, and 4 $HER2^+$ cell lines (Supplemental Table S2). Transcriptomic findings were also validated on 61 RNA tumor samples extracted from patient-derived xenografts (PDXs) derived from 41 TNBCs, 15 ER^+ carcinomas, and 5 $HER2^+$ carcinomas. No PDXs were derived from MBC.

RNA and DNA Extractions

RNA was extracted using the RNeasy kit (Qiagen, Courtaboeuf, France), followed by the RNA cleanup kit (Macherey-Nagel, Hoerdt, France). The quality of each RNA sample was measured with an Agilent 2100 bioanalyzer (RNA integrity number >7), and the quantity of RNA was measured by spectrophotometry at 260 nm. For the transcriptomic analysis validation set, comprising 526 breast tumors, cell lines, and PDXs, total RNA was extracted from samples using acid-phenol guanidium as previously described.[8] RNA quality was determined by electrophoresis on agarose gels, ethidium bromide staining, and visualization of the 18S and 28S RNA bands in UV light. DNA extraction and preparation for microarray experiments were performed by the Institut Curie Biological Resource Center. Before DNA isolation, a tissue section of tumor fragments was sampled and stained with hematoxylin and eosin to evaluate tumor cellularity. DNA was extracted from frozen tumor samples using a standard phenol/chloroform procedure. The quality of DNA was assessed on agarose gel. When a smear was observed instead of a band, the sample was discarded.

Gene Expression Analysis

Gene expression analysis was performed on 19 MBCs and 35 of 36 non-MBC BLCs. The RNA microarray used in this study was the GeneChip Human Genome U133 Plus 2.0 Array (Affymetrix, Santa Clara, CA), containing 54,613 probe sets. The transcriptomic data discussed in the present article, in addition to the genomic data from Affymetrix Genome Wide Human SNP Array 6.0 , have been deposited in the National Center for Biotechnology Information Gene Expression Omnibus (*https://www.ncbi.nlm.nih.gov/geo*; accession number GSE114269). Microarray data were simultaneously normalized using the GC robust multiarray average package 1.2 in the R environment (R Development Core Team). In this study, expression data were used to determine correlations between the genomic status of genes and their level of expression by comparing the log2 expression signal and the DNA copy number signal using Pearson's correlation test. A correlation between expression levels with respect to DNA copy number was considered significant when $R > 0.4$ and $P < 0.05$. Supervised analysis was performed to measure differential gene expression between MBC and non-MBC BLCs using a two-sided Welch t-test with $P < 0.05$ fold change >3. Partek GS software version 6.5 build 6.10.1020 (Partek Inc., St. Louis, MO) was used to generate unsupervised hierarchical clustering of expression data (Ward's method with Euclidean distance). Mean linkage was based on Pearson's dissimilarity of the 679

microarray probe sets with an interquartile range >3. A gene ontology–based determination of the biological pathways implicating differentially expressed genes in MBC and non-MBC BLC was performed using PANTHER pathway classification system version 13.0 (GENEONTOLOGY Unifying Biology, *http://www.pantherdb.org*, last accessed April 30, 2018), with the use of the Bonferroni correction for multiple testing.[9]

RT-qPCR for *BCLG* Expression

BCLG expression level was assessed in 19 MBCs, 41 non-MBC BLCs (34 from the initial cohort of 36 cases and seven additional samples collected previously with RNA available), 526 breast cancers corresponding to all molecular subtypes, 38 cell lines, and 61 PDXs. Quantitative values were obtained from the cycle number (C_T value) at which the increase in the fluorescence signal associated with exponential growth of PCR products started to be detected by the laser detector of the ABI Prism 7900 sequence detection system (Perkin-Elmer Applied Biosystems, Foster City, CA), using Perkin-Elmer Applied Biosystems analysis according to the manufacturer's manuals. Because the precise amount of total RNA added to each reaction mix (based on optical density) and its quality (ie, lack of extensive degradation) are difficult to assess, we also quantified transcripts of the *TBP* gene (GenBank accession number NM_003194) encoding the TATA box-binding protein (a component of the DNA-binding protein complex TFIID) as an endogenous RNA control and normalized each sample on the basis of its *TBP* content.[10] Results, presented as N-fold differences in *BCLG* gene expression relative to the *TBP* gene and termed NBCLG, were determined as NBCLG $= 2^{\Delta CT}$ sample, where the ΔC_T value of the sample was determined by subtracting the C_T value of the *BCLG* gene from the C_T value of the *TBP* gene. The NBCLG values of the human tumor samples were subsequently normalized so that the median of the NBCLG values for the 16 normal breast tissues was equal to 1. The results were represented as follows: transcript level $\geq$3 was considered to indicate overexpression, and transcript level $\leq$0.33 was considered to indicate underexpression. Because no normal equivalent counterpart was available, the NBCLG values of the cell lines and PDX samples were subsequently normalized considering 1 to be the lowest quantifiable transcript level ($C_T = 35$). The primers for *TBP* and *BCLG* (human and/or murine, long and short isoforms for *BCLG*) were chosen with the assistance of the Oligo 6.0 program (National Biosciences, Plymouth, MN). Overexpression of *BCLG* was defined as a *BCLG* expression level (NBCLG) >4. The dbEST and nr databases were scanned to confirm the total gene specificity of the nucleotide sequence chosen for the primers and the absence of single-nucleotide polymorphisms (SNPs). The nucleotide sequences of the oligonucleotide hybridization primers are given in Table 1. To avoid amplification of contaminating genomic DNA, one of the two primers was placed at the junction between two exons or on two different exons. Agarose gel electrophoresis was used to verify the specificity of PCR amplicons. The conditions of cDNA synthesis and PCR were as previously described.[11] *BCLG* gene transcription leads to three different transcripts. Two of them, differing from each other in terms of 5′-UTR configuration, led to the same protein after translation, the long variant of BCLG (BCLG-L). The third transcript encodes for the short variant of BCLG (BCLG-S). To document which of the variants is expressed in breast cancer, the *BCLG* gene expression level was compared in 19 MBCs and 21 breast tumors from the cohort of 526 samples considered to be high expressers, using primer pairs specific to long and short variants of BCLG in a RT-qPCR assay.

Genomic Alteration Analysis

SNP mapping assays were performed according to the manufacturer's protocol (Affymetrix, Santa Clara, CA). Briefly, 250 ng of genomic DNA from 19 MBCs and 36 non-MBC BLCs were digested with both *Nsp* and *Sty* restriction enzymes in independent parallel reactions, ligated to the adaptors, and amplified by PCR using a universal primer. After purification of PCR products with SNP clean magnetic beads (Agencourt Bioscience, Beverly, MA), amplicons were quantified, fragmented, labeled, and hybridized to Affymetrix Genome-Wide Human SNP 6.0 array. Targets were prepared when 45 μg of amplified DNA was available and when the target size was situated between 250 and 2000 bp and were then hybridized according to the manufacturer's recommendations. After washing and staining, the arrays were scanned to generate .cel files for downstream analysis. Normalization was performed using a Genotyping console (GenomeWideSNP_6.hapmap270.na31.r1.a5.ref) provided by Affymetrix (GTC3.0.1). Genomic alterations were evaluated according to the Genome Alteration Print (GAP) method.[12] Copy number and allelic content profiles were

Table 1 Nucleotide Sequences of the Quantitative RT-PCR Primers for *BCLG*

Transcript	Forward	Reverse
TBP	5′-TGCACAGGAGCCAAGAGTGAA-3′	5′-CACATCACAGCTCCCCACCA-3′
BCLG human	5′-GCCAGCAGTGGTCCAGGTGTCT-3′	5′-CTACTCGGTTGGCAATGGAAATGA-3′
BCLG long	5′-CACGTGCCTGTAGCTTCAAGTTC-3′	5′-AAGCCTTATCTTTCTTCAGCTTTCTT-3′
BCLG short	5′-CACGTGCCTGTAGCTTCAAGTTCT-3′	5′-GGGATGAAGGCAGTGTCCTTTCT-3′
BCLG murin	5′-AGCACTGGCCCAGGTCTCTGA-3′	5′-CAGCCACTCTGTTGGCAATACAAG-3′

detected for each tumor based on the overall pattern of alterations, as previously described and validated.[12] The cutoffs for alteration events (gains, losses, focal amplifications) were adapted according to the inferred ploidy. For near-diploid tumors, the genomic region with inferred copy number ≤ 1 or ≥ 3 and ≥ 6 were considered to be regions of loss, gain, and focal amplification, respectively. For near-tetraploid tumors, the copy number cutoffs used to define regions of loss, gain, and focal amplification were 2, 6, and 8, respectively. The minimum regions of amplification covering at least 25 consecutive SNPs with the same copy number status were considered to be recurrent regions when the frequency of alterations was >15% in the tumor samples tested, after exclusion of genomic variants according to the genomic variants database. Tumor profiles were visualized with GAP software (Institut Curie).[12]

Affymetrix Genome-Wide Human SNP Array 6.0 genomic data allowed the assessment of large-scale state transition (LST) status as described previously.[13] Briefly, the LST score is the number of breakpoints between regions longer than 10 Mb after filtering out regions shorter than 3 Mb. Different cutoffs for the LST score were introduced for near-diploid and near-tetraploid tumors to separate *BRCA1/2* intact and deficient samples, regardless of the inactivation process. This process resulted in segregation of samples into high LST and low LST samples according to the LST score. The LST status was able to be assessed in 16 of 19 MBCs and 34 of 36 non-MBC BLCs.

Methylation of the *BRCA1* promoter was assessed by methyl-specific PCR after bisulfite conversion using the MethylDetector Kit (Active Motif, Carlsbad, CA), as previously described.[14] Somatic and germline *BRCA1/2* mutational status was determined as previously described.[13] Primer sequences designed for *BRCA1/2* mutational status and methylation of the *BRCA1* promoter assessment are detailed in Tables 2, 3, and 4, respectively.

IHC Analysis

Tissue sections (4 μm thick) were cut from formalin-fixed, paraffin-embedded representative blocks of 16 of 19 MBCs and 27 of 36 non-MBC BLCs sampled and gathered in a tissue microarray. After deparaffinization according to standard procedures, all sections were subjected to heat-induced antigen retrieval in citrate buffer (pH 6.1). Antibody against BCLG (clone ab184925, Abcam, Cambridge, UK)

Table 2 Nucleotide Sequences of the PCR Primers for *BRCA1* Mutational Status Assessment

BRCA1 amplicons	Screened region	Forward	Reverse
02speA, 02speB	UTR-5′-170—IVS02+127	5′-AATAAAGGACGTTGTCATTAGTTCT-3′	5′-AGCAATTACAATAGCCTAATCTTAC-3′
03eA, 03eB	IVS02-164—IVS03+201	5′-TTCCTGACACAGCAGACATTT-3′	5′-ATGTCAAAACTTTACCAGGAAC-3′
05eA, 05eB	IVS04-128—IVS05+110	5′-GCCATTACTTTTTAAATGGCTC-3′	5′-TTATAAATTTTTCTGATGAATGGTT-3′
06eA, 06eB	IVS05-121—IVS06+122	5′-AGAGGTTTTCTACTGTTGCTG-3′	5′-CAGAACTAAAATTAACCTAGACT-3′
07eA, 07eB	IVS06-125—IVS07+114	5′-GGTAACCTTAATGCATTGTCTT-3′	5′-AAGGCAGGAGGACTGCTTCT-3′
08eA, 08eB	IVS07-154—IVS08+108	5′-TGTGTAAATTCCTGGGCATT-3′	5′-CAAAGCTGCCTACCACAAATA-3′
09eA, 09eB	IVS08-121—IVS09+132	5′-CCACAGTAGATGCTCAGTAAA-3′	5′-AACAAACTGCACATACATCCC-3′
10eA, 10eB	IVS09-121—IVS10+109	5′-CTAAATAAGATTGGTCAGCTTTCT-3′	5′-TTTTGTGGGTTGTAAAGGTCC-3′
11-01eA, 11-01eB	IVS10-125—c.1097	5′-GTTTATGAGGTTAGTTTCTCTAA-3′	5′-TCTCTAGGATTCTCTGAGCAT-3′
11-02eA, 11-02eB	c.942—c.1463	5′-AAGGAGCCAACATAACAGATG-3′	5′-GTAACAAATGCTCCTATAATTAG-3′
11-03eA, 11-03eB	c.1326—c.1850	5′-TAAAAGTGAAAGAGTTCACTCC-3′	5′-GTAGAAGACTTCCTCCTCAG-3′
11-04eA, 11-04eB	c.1702—c.2222	5′-CCTAACCCAATAGAATCACTC-3′	5′-GACACTTTAACTGTTTCTAGTTT-3′
11-05eA, 11-05eB	c.2084—c.2600	5′-ATACTTTCCCAGAGCTGAAGT-3′	5′-TGGCGCTTTGAAACCTTGAAT-3′
11-06eA, 11-06eB	c.2440—c.2959	5′-CTAATTCATGGTTGTTCCAAAG-3′	5′-TGATGGGAAAAAGTGGTGGTA-3′
11-07eA, 11-07eB	c.2799—c.3238	5′-TCAGAAAGATAAGCCAGTTGAT-3′	5′-ATTTTGGCCCTCTGTTTCTAC-3′
11-08eA, 11-08eB	c.3077—c.3629	5′-TTAGCCGTAATAACATTAGAGAA-3′	5′-TCTAATTTCTTGGCCCCTCTT-3′
11-09eA, 11-09eB	c.3468—c.3924	5′-TGGTGAAATAAAGGAAGATACTA-3′	5′-TGCAGTCAAGTCTTCCAATTC-3′
11-10eA, 11-10eB	c.3789—IVS11+127	5′-GAAGAATAGCTTAAATGACTGC-3′	5′-TTCAAGTTTAAGAAGCAGTTCC-3′
12eA, 12eB	IVS11-110—IVS12+109	5′-TGTGTGACATGAAAGTAAATCC-3′	5′-CCATTAATTCAAAGAGATGATGT-3′
13eA, 13eB	IVS12-123—IVS13+110	5′-TTGTAGTTCCATACTAGGTGAT-3′	5′-CTGAGCAAGGATCATAAAATGT-3′
14eA, 14eB	IVS13-137—IVS14+112	5′-TCTGCCTGATATACTTGTTTAAA-3′	5′-AATGCCTGTATGCAAAAAACTG-3′
15eA, 15eB	IVS14-122—IVS15+134	5′-TGCCAGTCATTTCTGATCTCT-3′	5′-GTGGGCTTAATTAAGTATAACA-3′
16eA, 16eB	IVS15-122—IVS16+116	5′-ATTCATGTACCCATTTTTCTCTT-3′	5′-GTGATTGTTTTCTAGATTTCTTC-3′
17eA, 17eB	IVS16-131—IVS17+133	5′-ATAGTTCCAGGACACGTGTA-3′	5′-CGATCTCCTAATCTCGTG-3′
18eA, 18eB	IVS17-115—IVS18+107	5′-ATAAATCCAGATTGATCTTGG-3′	5′-GTAACTCAGACTCAGCATCA-3′
19eA, 19eB	IVS18-119—IVS19+106	5′-AAGGACCTCTCCTCTGTCAT-3′	5′-TGTGCATTGTTAAGGAAAGTG-3′
20eA, 20eB	IVS19-205—IVS20+151	5′-ATGAGGTTTCACCATGTTGGT-3′	5′-GAAGAGTGAAAAAAGAACCTGT-3′
21eA, 21eB	IVS20-116—IVS21+113	5′-AAGAAAAGCTCTTCCTTTTTGAA-3′	5′-TCTAGAACATTTCAGCAATCTG-3′
22eA, 22eB	IVS21-183—IVS22+174	5′-GTGGCAAATTGACTTAAAATCC-3′	5′-CAGTTCTCAAATCCTTACCCA-3′
23eA, 23eB	IVS22-229—IVS23+198	5′-AGGGGTGGTGGTACGTGTCT-3′	5′-CCATGGAAACAGTTCATGTATT-3′
24eA, 24eB	IVS23-120—UTR-3′+108	5′-TTAGCTTCTACCTCATTAATCC-3′	5′-AGGACAGTAGAAGGACTGA-3′

Table 3 Nucleotide Sequences of the PCR Primers for *BRCA2* Mutational Status Assessment

BRCA2 amplicons	Screened region	Forward	Reverse
	UTR-5′-156—IVS02+119	5′-AGGAGATGGGACTGAATTAGA-3′	5′-CACATAAGGAACAGTTTATGG-3′
03eA 03eB	IVS02-118—IVS03+123	5′-CAAAAGTAATCCATAGTCAAGAT-3′	5′-AGAGGCCAGAGAGACTGATT-3′
04eA 04eB	IVS03-121—IVS04+120	5′-AACTCCCTATACATTCTCATTC-3′	5′-AGATCTTCTACCAGGCTCTTA-3′
05eA 05eB	IVS04-121—Exon_6 (c.501)	5′-AATATCTAAAAGTAGTATTCCAAC-3′	5′-TGTATGAAACAAACTCCCAC-3′
06eA 06eB	Exon_5 (c.447)—IVS06+118	5′-ACATGTAACACCACAAAGAGAT-3′	5′-ATTGCCTGTATGAGGCAGAAT-3′
07eA 07eB	IVS06-119—IVS07+122	5′-ATTCTGCCTCATACAGGCAAT-3′	5′-CACACTTATCAAAGACATTATCT-3′
08eA 08eB	IVS07-114—IVS08+124	5′-TGTTTCAAATGTGTCATGTAATC-3′	5′-GACTTTCTCAAAGGCTTAGATA-3′
09eA 09eB	IVS08-121—IVS09+114	5′-GACCTAGGTTGATTGCAGATA-3′	5′-AGAGGTTGCGGTAAACCGAG-3′
10-01eA 10-01eB	IVS09-120—c.1135	5′-TACTGATATGTAATATTTAGCACA-3′	5′-CACTCTCAAAGGGCTTCTGA-3′
10-02eA 10-02eB	c.996—c.1457	5′-TTTCCATGAAGCAAACGCTGA-3′	5′-TGCTTTACTGCAAGAATGCAG-3′
10-03eA 10-03eB	c.1320—c.1749	5′-TACTTCAGAGAATTCTTTGCCA-3′	5′-CAAAGTGGATATTAAACCTGCA-3′
10-04eA 10-04eB	c.1610—IVS10+126	5′-AAAGTGGACTGGAAATACATAC-3′	5′-GTATACAGATGATGCCTAAGAT-3′
11-01ieA 11-01ieB	IVS10-97—c.2226	5′-ACTGTGCCCAAACACTAC-3′	5′-TTGTACTGGGTGACATGC-3′
11-01eA 11-01eB	c.2040—c.2333	5′-AGTAATCTCTCAGGATCTTGAT-3′	5′-ACATCCTTGGAAGTAGGAGTT-3′
11-02eA 11-02eB	c.2186—c.2715	5′-TAAAAGAAGAGGTCTTGGCTG-3′	5′-ATTTCCTAAAGCAAGATTATTCC-3′
11-03eA 11-03eB	c.2552—c.3081	5′-TCAACCAAAACACAAATCTAAGA-3′	5′-GCTCTTCTTAATGTTATGTTCAG-3′
11-04eA 11-04eB	c.2918—c.3445	5′-CGGACATCTCCTTGAATATAG-3′	5′-TCTGGTTTTCAGGCACTTCAA-3′
11-05eA 11-05eB	c.3306—c.3814	5′-TTTAACACCTAGCCAAAAGGC-3′	5′-TTGAAACAACAGAATCATGAC-3′
11-06eA 11-06eB	c.3671—c.4199	5′-GCACAAAACTGAATGTTTCTAC-3′	5′-TGACATGCTTCTTGAGCTTTC-3′
11-07eA 11-07eB	c.4061—c.4584	5′-CGGACTTGCTATTTACTGATC-3′	5′-GCTAGCTGTATGAAAACCCAA-3′
11-08eA 11-08eB	c.4443—c.4969	5′-GGAAACAGACATAGTTAAACAC-3′	5′-TTGTGTAACAAGTTGCAGGAC-3′
11-09eA 11-09eB	c.4834—c.5376	5′-CCACCTAAGCTCTTAAGTGAT-3′	5′-GGATATTACTTTGGAAAAACTAG-3′
11-10eA 11-10eB	c.5209—c.5739	5′-GATACTTATTTAAGTAACAGTAGC-3′	5′-ACATTCATCATTATCTAGAGAG-3′
11-11eA 11-11eB	c.5599—c.6127	5′-ACAGACAGTTTCAGTAAAGTAA-3′	5′-TTTGGGATATTAAATGTTCTGG-3′
11-12eA 11-12eB	c.5985—c.6512	5′-CGCAAGACAAGTGTTTTCTGA-3′	5′-ACTTTGGTTCCTAATACCAACT-3′
11-13eA 11-13eB	c.6349—c.6832	5′-TGTGTAAACTCAGAAATGGAAAA-3′	5′-TAAGGGGCTCTCCTCTTCTT-3′
11-14eA 11-14eB	c.6668—IVS11+136	5′-TTGAAACAGAAGCAGTAGAAATT-3′	5′-TCCCCCAAACTGACTACACA-3′
12eA 12eB	IVS11-141—IVS12+139	5′-TAGGTCACTATTTGTTGTAAGTA-3′	5′-TAAAGAGGTCCTTGATTAGGC-3′
13eA 13eB	IVS12-121—IVS13+119	5′-CTGTTACATTCACTGAAAATTG-3′	5′-TAAAACGGGAAGTGTTAACTTC-3′
14-01eA 14-01eB	IVS13-120—c.7207	5′-ACAAAACAGTTACCAGAATAGTA-3′	5′-TTGGTCTGCCTGTAGTAATCA-3′
14-02eA 14-02eB	c.7133—IVS14+16	5′-CAGGACATCCATTTTATCAAGT-3′	5′-AATTGTCATACAATACCTAAAGG-3′
14-03eA 14-03eB	c.7313—IVS14+153	5′-ATGGACATGGCTCTGATGATA-3′	5′-TTAAACCTAATCTTTGGATTTAGA-3′
15eA 15eB	IVS14-139—IVS15+120	5′-TGAACTCCCGACCTCAGAT-3′	5′-ATTCATCCATTCCTGCACTA-3′
16eA 16eB	IVS15-138—IVS16+144	5′-TGTTTTTGTAGTGAAGATTCTAG-3′	5′-TGCTTAACCATAATGCACTTAAAA-3′
17eA 17eB	IVS16-122—IVS17+119	5′-GAACTCATAAAAACTTAATGATCT-3′	5′-GATGGCAACTGTCACTGACAA-3′
18-01eA 18-01eB	IVS17-150—c.8222	5′-GAAACAATATATTCCTAGCTACA-3′	5′-TTTAAGACAGCTAAGAGGGGA-3′
18-02eA 18-02eB	c.8071—IVS18+215	5′-TCTGACATAATTTCATTGAGCG-3′	5′-TGGAAATGCATTATTTAAGCTCA-3′
19-01eA 19-01eB	IVS18-127—c.8467	5′-CTTCCTAAGACTTTTTAAAGTGA-3′	5′-GAATAATTACATCAACACAACCA-3′
19-02eA 19-02eB	c.8332-24—IVS19+131	5′-CAATATATTTATTAATTTGTCCAG-3′	5′-CTGCAGTGAACCAAGATCAC-3′
20eA 20eB	IVS19-118—IVS20+120	5′-TAATCTCAGCCTCCCAAAGTT-3′	5′-TAAAGTCAATTTACTACTCAA-3′
21eA 21eB	IVS20-121—IVS21+120	5′-GCAGTTATATAGTTTCTTATCTTTA-3′	5′-ATCCCTTTTGAGAAATGCAGC-3′
22eA 22eB	IVS21-119—IVS22+122	5′-ACACCCTTAAGATGAGCTCTA-3′	5′-GTGGATTTTGCTTCTCTGATAT-3′
23eA 23eB	IVS22-120—exon_24 (c.9152)	5′-AAATCCACTACTAATGCCCAC-3′	5′-GGCTGGTAAATCTGAAATAAAAT-3′
24-01eA 24-01eB	Exon_23 (c.9094) — IVS24+17	5′-AAAACTCAGTATCAACAACTACC-3′	5′-ACTATATTGTGCATTACCTGTTT-3′
24-02eA 24-02eB	c.9118-7—IVS24+141	5′-TCTGTAGGTTTCAGATGAAATTTT-3′	5′-GAGGTTCAAAGAGGCTTACTT-3′
25eA 25eB	IVS24-82—IVS25+116	5′-CATATTAGAGTTTCCTTTCTTGC-3′	5′-CTTTACCTCACATACTACCTCA-3′
26eA 26eB	IVS25-120—IVS26+144	5′-AGGGTTTTTCATTCTTTTTTGGT-3′	5′-AACTATACTTACAGGAGCCAC-3′
27-01eA 27-01eB	IVS26-117—c.10007	5′-GAGGGAGACTGTGTGTAATAT-3′	5′-CTTTCCAAAAGAGAAATTTCATTG-3′
27-02eA 27-02eB	c.9877—UTR-3′+134	5′-GCTGCACAGAAGGCATTTCA-3′	5′-CTTTGCTCATTGTGCAACATAA-3′

was then applied and staining was performed with the Vectastain Elite ABC peroxidase mouse IgG kit (Vector Laboratories, Burlingame, CA), with diaminobenzidine (Dako, Santa Clara, CA) as chromogen. Semiquantitative assessment of expression was performed using a light microscope by two independent readers (P.R., G.D.) according to the following algorithm: the staining intensity in tumor cells was graded on a scale of 0 (no expression) to 3 (for highest intensity). The entire section was assessed. The percentage of positive tumor cells was calculated for each

Table 4 Nucleotide Sequences of the *BRCA1* Promoter Methylation Status Assessment

BRCA1 promoter status	Forward	Reverse
BRCA1 promoter, methylated reaction	5′-TTGGTTTTTGTGGTAATGGAAAAGTGT-3′	5′-CAAAAAATCTCAACAAACTCACACCA-3′
BRCA1 promoter, unmethylated reaction	5′-TCGTGGTAACGGAAAAGCGC-3′	5′-AAATCTCAACGAACTCACGCCG-3′

specimen, and this percentage score was multiplied by the staining intensity to obtain a final semiquantitative H score that ranged from 0 to 300. The mean H score based on scoring by the two readers was attributed to each tumor sample. The *t*-test was used to assess specificity of IHC staining between the two groups. Spearman's rank correlation test was used to assess the correlation between H score and BCLG expression level.

RNA Sequencing

Library preparation and paired-end (2 × 100 bp) RNA sequencing were performed on six MBCs by IntegraGen (Evry, France) using a TruSeq RNA sample preparation kit and HiSeq 2000 platform (both from Illumina, San Diego, CA), respectively. A mean of 130 million reads were obtained for each sample. RNA sequencing raw reads were mapped using TopHat version 2.0.6 (Johns Hopkins University Center for Computational Biology, Baltimore, MD; *https://ccb.jhu.edu/software/tophat/index.shtml*, last accessed April 9, 2018) and Bowtie version 2.0.4 (Johns Hopkins University Center for Computational Biology; *http://bowtie-bio.sourceforge.net/bowtie2/index.shtml*) against UCSC hg19 genome (National Center for Biotechnology Information build 37.1). Single-nucleotide variants and transcriptome quantitative analysis were performed using SAM tools version 0.1.8 and Cufflinks version 2.0.2 (Cole Trapnell, Seattle, WA; *http://cole-trapnell-lab.github.io/cufflinks*, last accessed April 9, 2018), respectively. Expression levels of validated mutations were determined using RNA sequencing data. Gene fusion analyses were performed using two validated tools: TopHat-fusion version 2.0.4 and deFuse version 0.6[15–17] (Supplemental Table S3). Gene fusion validation was performed using RT-qPCR and Sanger sequencing (primer sequences used are detailed in Table 5).

TNBC Subtyping

Gene expression levels obtained from GeneChip Human Genome U133 Plus 2.0 Array analysis were submitted as preprocessed and normalized data to the web-based subtyping tool TNBCtype for candidate TNBC samples according to the web-interface developers' instructions (*http://cbc.mc.vanderbilt.edu/tnbc*, last accessed April 9, 2018). Briefly, gene expression values expressed in cnv format (ASCII format file delimited by comma) were uploaded for each sample in the genome-wide TNBC gene expression matrix based on a data set of 3247 gene expression profiles from 21 breast cancers. After format checking, 18 of 19 MBC cases and 28 of 35 non-MBC BLC cases were classified into one of the six predicted TNBC subtypes based on the gene expression matrix profiles, with corresponding correlation coefficients and permutation *P* values. TNBC subtypes include basal-like 1 (BL1) and basal-like 2 (BL2) subtypes, immunomodulatory (IM) subtype, mesenchymal (M) subtype, mesenchymal stem-like (MSL) subtype, and luminal androgen (LAR) subtype, defined according to shared unique profiles among 587 TNBC samples. Tumors with no predominant gene expression profile were classified as unstable.[18] Finally, in accordance with the refined classification of TNBCtype, the so-called TNBCtype-4,[19] IM and MSL cases were reassessed to the second highest correlated centroid among BL1, BL2, M, and LAR subtypes.

Survival Analysis

Survival analysis was conducted on an extended cohort of 32 MBCs and 44 non-MBC BLCs. Differences between groups were compared by a χ^2 test or Fisher's exact test for categorical variables. Disease-free survival was defined as the time elapsed between diagnosis of the primary tumor and occurrence of the first event, either recurrence,

Table 5 Nucleotide Sequences of the Gene Fusion Validation Study

Transcript	Forward	Reverse
KIAA1467-ETV6	5′-CCAGCCTTCACCACCTTTAC-3′	5′-GCGAAAGTCCTCTTTGGTCA-3′
STAG1-CEP70	5′-CCTCTCCAGCAATGATTACTTC-3′	5′-CAATGCGATCTTCTTCCTCCT-3′
CCDC132-AKAP9	5′-TTTCAAGAATTACCATAGGAACACG-3′	5′-TTGAGCTGCTCTATTTCTTCTTCTC-3′
TNPO2-SHCBP1	5′-CGTAGCACTGGATCTGCTCA-3′	5′-CCATGGATATCTTGGGAATG-3′
IFT140-AATF	5′-AGACCCCTGCGGAGACTTTGT-3′	5′-TTTTTGTGGATTTTGCTTCG-3′
MAP3K11-RELA	5′-CCAGAATGGGAGGAGAAGGT-3′	5′-GAGCCTGGGGCAGGACTT-3′
ZFAND3-RNF8	5′-CCGAGCACCATGGGAGAC-3′	5′-CTCTGTTCAGCCAAACACCA-3′
TMEM123-MMP20	5′-CAGGTGCTAGCGCTGCTG-3′	5′-TCCACTTCTCAGCATTGTCG-3′
LOH12CR1-ETV6	5′-CAACAAATGCCAAATTGGAG-3′	5′-CGAGTCTTCCTCCATCCTGA-3′
CTSC-RAB38	5′-CAAACCTGCACCACTGACTG-3′	5′-CACTTTGCCACTGCTTCAAA-3′

metastasis, or death. Overall survival (OS) was based on a similar definition without taking the recurrence and metastasis events into account. Patients who were still alive and those free of disease recurrence at the time of last follow-up were censored at the date of last known contact. Both OS and disease-free survival (DFS) were estimated using the Kaplan-Meier method, and groups were compared using the log-rank test. Statistical significance was set at $P < 0.05$, and analyses were performed using R version 3.0.1 statistical software (The Comprehensive R Archive Network, Vienna, Austria; *http://cran.r-project.org*, last accessed April 9, 2018).

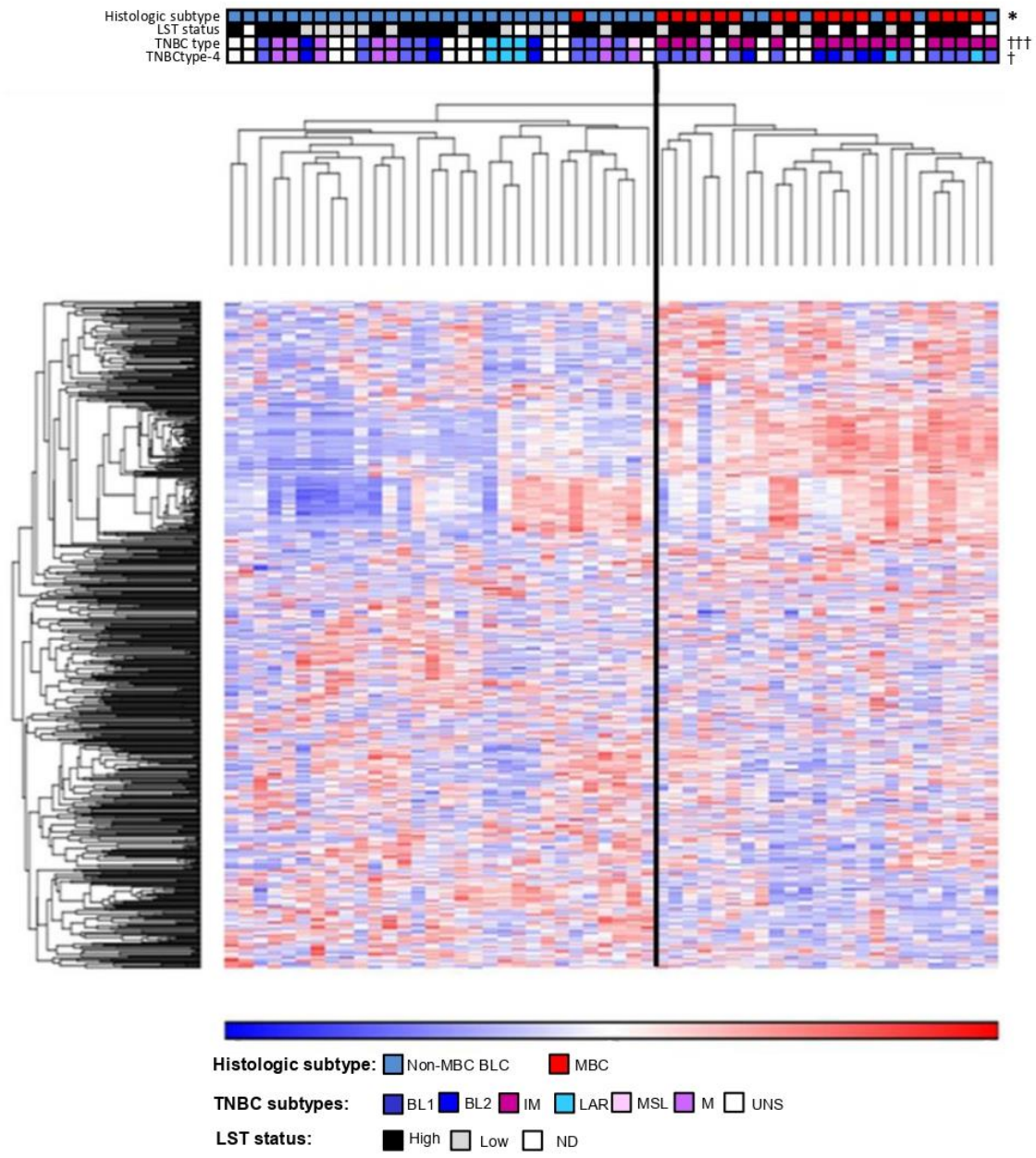


Figure 1 Transcriptomic analysis. Hierarchical clustering performed using Average Linkage based on Pearson's dissimilarity of the 679 microarray probe set with an interquartile range >3. Blue signal indicates low expression; red signal, high expression. Histological subtype (**first row**); large-scale state transition (LST) status of tumor sample (**second row**); triple-negative breast cancer (TNBC) type according to Lehmann et al.[18] (**third row**); TNBCtype-4 according to Lehmann et al.[19] (**fourth row**). *$P < 0.05$ non-MBC BLC cases in cluster 1 versus versus composition in MBC cases in cluster 2; †$P < 0.05$, †††$P < 0.001$ subtype attribution in MBC versus non MBC BLC. BL1, basal-like 1; BL2, basal-like 2; IM, immunomodulatory; LAR, luminal androgen; M, mesenchymal; MSL, mesenchymal stem-like; ND, not determined; UNS, unstable.

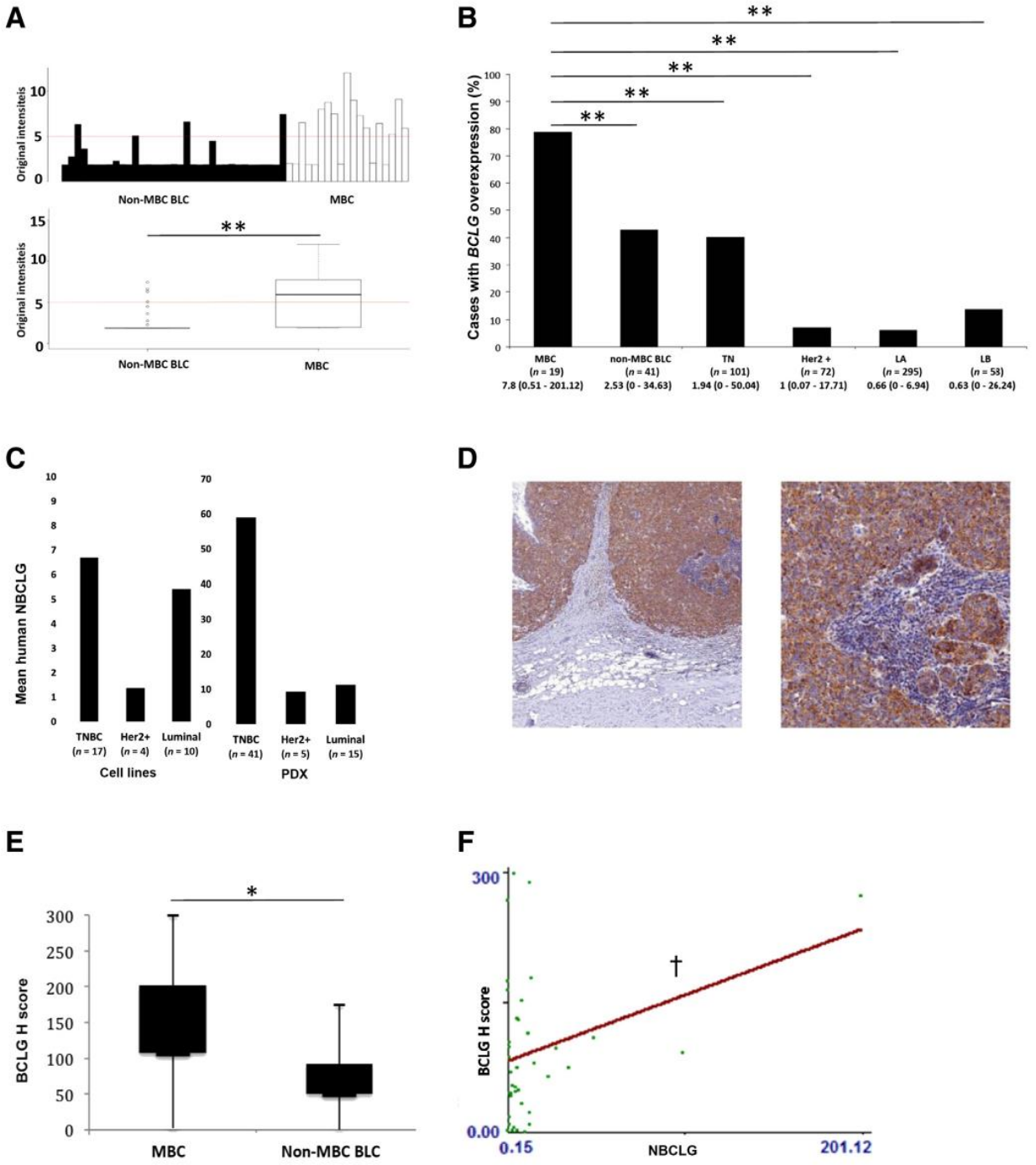


Figure 2 *BCLG* overexpression in medullary breast carcinoma (MBC). **A:** Affymetrix U133 Plus 2.0 analysis: *BCLG* expression level (NBCLG) bar plots in 19 MBCs and 35 non-MBC basal-like carcinomas (BLCs) and corresponding box plots. Red line indicates differential gene expression threshold as defined by the Welch *t*-test. **B:** Quantitative RT-PCR validation study: percentage of cases with *BCLG* overexpression in 19 MBCs, 41 BLCs, and 526 invasive breast carcinomas from different molecular classes. The median NBCLG (range) is given for each group. **C:** Information on epithelial expression of *BCLG*: mean human NBCLG in 31 tumor cell lines (and seven nontumor cell lines) (data not shown) and in 61 patient-derived xenografts (PDXs) according to molecular classes. **D:** *BCLG* expression by immunohistochemistry in epithelial cancer cells: *BCLG* staining on MBC formalin-fixed, paraffin-embedded sections. **E:** *BCLG* expression by immunohistochemistry in epithelial cancer cells: median BCLG H scores box plots (clone ab184925, Abcam) in 16 MBCs and 27 non-MBC BLCs. **F:** Spearman linear correlation between *BCLG* expression level and BCLG H score in MBC and non-MBC BLC. Red line indicates linear regression line ($r = 0.328$). $^{*}P < 0.05$, $^{**}P < 0.01$; $^{\dagger}P < 0.05$ (Spearmann's rank correlation). Original magnification: ×25 (**D**, **right panel**); ×400 (**D, left panel**).

Table 6 Copy Number Alterations of *BCLG* Chromosomal Region (12p13.2) in Medullary Breast Carcinoma (MBC) and Non-MBC Basal-Like Carcinoma (BLC)

Genomic status	MBC, *n* (%)	Non-MBC BLC, *n* (%)
Gain	8 (42)	25 (69)
Focal amplification	2 (11)	9 (25)
Loss	6 (32)	5 (14)

$n = 19$ MBC; $n = 36$ non-MBC BLC.

Results

MBC Has a Distinct Gene Expression Profile from Non-MBC BLC

Among the studied series, two different transcriptomic groups of tumors were identified, as shown in dendro-heatmaps in Figure 1 (unsupervised analysis). The first cluster was predominantly [29 of 30 (97%)] composed of non-MBC BLCs, and the second cluster was mostly [18 of 24 (75%)] composed of MBCs ($P < 0.05$). Supervised transcriptomic data were then analyzed to precisely identify deregulated genes in these two groups and showed that 154 genes were underexpressed in MBCs compared with non-MBC BLCs and that 92 genes were overexpressed in MBCs compared with non-MBC BLCs (Supplemental Table S4). *BCLG* (alias *BCL2L14*) was the most differentially overexpressed gene in MBCs compared with non-MBC BLCs (fold change = 23) (Figure 2A). Interestingly, those non-MBC BLCs with *BCLG* overexpression were associated with a high TILs infiltrate (Supplemental Table S5).

A Gene Ontology study was performed to identify the biological processes in which these genes are involved. Significantly represented Gene Ontology pathways identified are listed in Supplemental Table S6 (supervised analysis). On the basis of the overexpressed genes in MBC, 110 Gene Ontology biological processes were identified with most immunity and inflammatory response–related pathways, whereas six pathways related to developmental processes resulted from the analysis based on underexpressed genes in MBC.

BCLG Overexpression in MBC Is Confirmed by RT-qPCR in the Validation Cohort

This validation analysis found that the NBCLG ranged from 0.18 to 2.82 among normal breast samples and that *BCLG* was overexpressed (transcript level ≥3 compared with the median NBCLG of normal tissue normalized as 1) in 78.9% of MBCs versus 42.9% of non-MBC BLCs and 40.2% of TNBCs ($P = 0.006$). Among the 41 of the 101 TNBCs that overexpressed *BCLG*, three were defined as atypical MBC and one was defined as MBC. *BCLG* expression was higher in TNBC than in the other subgroups (Figure 2B). It was therefore confirmed that *BCLG* overexpression is significantly more frequently found in MBC.

Analysis of overexpression of the three *BCLG* transcripts (*Materials and Methods*) in MBC revealed that the median NBCLG-L was 31.68 (range, 0.00 to 2372.84) and the median NBCLG-S was 1.11 (range, 0.00 to 47.57) in these tumors (data not shown). The long variant transcript of the gene, encoding for a 327–amino acid protein, was therefore mainly responsible for *BCLG* overexpression in breast tumors.

To determine whether *BCLG* overexpression in breast tumors was related to the epithelial or stromal component of the tumor and because of the high density of TILs in MBC, *BCLG* expression was assessed in epithelial cell lines and PDX tumors (Figure 2C). *BCLG* gene expression was null or minimal in seven nontumor cell lines. It is noteworthy that 7 of 31 tumor cell lines (22.6%) (HCC-1187, HCC-38, HCC-70, MDA-MB-468, HCC-1937, MDA-MB-415, BT474) had a median NBCLG >4. Five of these seven cell lines were derived from TNBC tumors. The highest *BCLG* expression was observed in the HCC-70 cell line

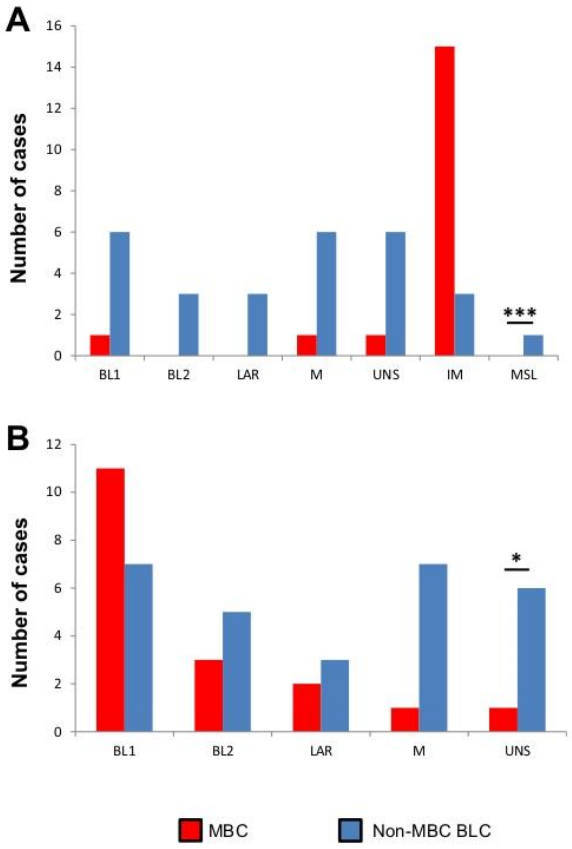


Figure 3 Medullary breast carcinoma (MBC) and non-MBC basal-like carcinoma (BLC) distribution within the transcriptomic classes defined by Lehmann et al.[18] **A:** Classification based on six classes. **B:** Classification based on four classes.[19] $*P < 0.05$, $***P < 0.001$. BL1, basal-like 1; BL2, basal-like 2; IM, immunomodulatory; LAR, luminal androgen; M, mesenchymal; MSL, mesenchymal stem-like; UNS, unstable.

($N = 62.7$) and the HCC 1187 cell line ($N = 20.8$), known to belong to the IM subgroup according to Lehmann et al.[18] Notably, the two ER^+ cell lines (MDA-MB-415 and BT474) that expressed higher levels of *BCLG* lack progesterone receptor expression (Supplemental Table S2).[20] Overexpression of human BCLG expression ($C_T < 32$) was observed in 38 of 61 PDXs (62.3%), and murine *BCLG* expression was observed in only 5 of 61 cases (8.1%). Equally interesting, 27 of 38 PDXs (69.2%) overexpressing human *BCLG* were TNBC PDXs. Therefore, it can be concluded that *BCLG* overexpression is related to the epithelial component of the tumor.

Finally, epithelial expression of *BCLG* was also confirmed by IHC studies performed on 16 MBCs and 27 non-MBC BLCs. A granular cytoplasmic *BCLG* staining pattern was observed in carcinomatous cells. TILs did not express *BCLG* (Figure 2D). The median H score in MBC was higher than in non-MBC BLCs (107.5 versus 50.0, $P = 0.019$) (Figure 2E). Furthermore, MBC tumors were enriched with high BCLG H score cases (H score >200) (Supplemental Table S7). BCLG transcript and protein expression levels were correlated in this series (Figure 2F).

BCLG Overerxpression in MBC Is Not Related to Copy Number Alterations

BCLG copy number localized at chromosome 12p13.2 was assessed in all samples of the initial cohort (Table 6). *BCLG* gains were observed in 8 of 19 (42%), including two focal amplifications, of MBCs and 25 of 36 (69%), including nine focal amplifications, of non-MBC BLCs. Conversely, *BCLG* losses were observed in 6 of 19 MBCs (32%) and 5 of 36 non-MBC BLCs (14%). These results therefore indicate that *BCLG* gene overexpression in MBC is not related to copy number alterations. Similarly, no correlation was observed between *BCLG* expression level and high LST or low LST status in MBCs or non-MBC BLCs (data not shown).

RNA Sequencing Identifies Putative Fusion Candidate Genes in MBC

RNA sequencing identified eight putative fusion candidate genes common to the two analysis tools in six MBCs, which were then validated by RT-qPCR and Sanger sequencing

Table 7 Specific Loss of Heterozygosity (LOH) in Medullary Breast Carcinoma (MBC) and Non-MBC Basal-Like Carcinoma (BLC)

Start position	End position	Chromosome	Cytoband	MBC, %	Non-MBC BLC, %
Specific LOH in MBC group ($n = 19$)					
24309700	41814174	2	p23.3-p21	43	15
50184394	50210408	2p	p16.3	42	11
71779115	71814760	2p	p13.2	42	8
89253701	91,689,680	2p	p11.2-p11.1	42	17
95341388	100168258	2q	q11.1-q11.2	45	17
107002769	114193894	2q	q12.2-q13	57	26
133038729	133114536	2q	q21.2	58	25
52696792	64462524	4q	q11-q13.1	62	30
70982244	71169616	4q	q13.3	63	33
131866907	139609545	4q	q28.3-q31.1	79	50
189263717	190930894	4q	q35.2	89	61
3301088	3370201	7p	p22.2	63	33
16078412	16913789	7p	p21.2-p21.1	63	33
18408294	18807219	7p	p21.1	63	33
55182494	55205930	7p	p11.2	42	17
107591914	107623496	7q	q31.1	42	17
22250879	22392626	14q	q11.2	84	54
20964485	26289925	20p	p11.23-p11.1	45	12
46494273	46889259	20q	q13.3	42	17
62896110	62912463	20q	q13.33	79	25
54491633	54699522	23p	p11.22	68	39
75248351	76230011	23q	q13.3-q21.1	78	47
84224371	96955143	23q	q21.1-q21.33	72	38
Specific LOH in non-MBC BLC group ($n = 36$)					
162270070	162314921	6q	q26	21	50
1496606	36856268	9p	p24.3-p13.2	19	51
97370455	100817655	11q	q22.1	26	56
114219050	134926754	11q	q23.2-q25	38	74
33612829	47067159	18q	q12.2-q21.1	22	53

Genomic positions are provided according to Human Genome 19 references. Specific regions of LOH listed in the table were observed in ≥40% of cases. Specificity was defined by a Fisher test $P \leq 0.05$.

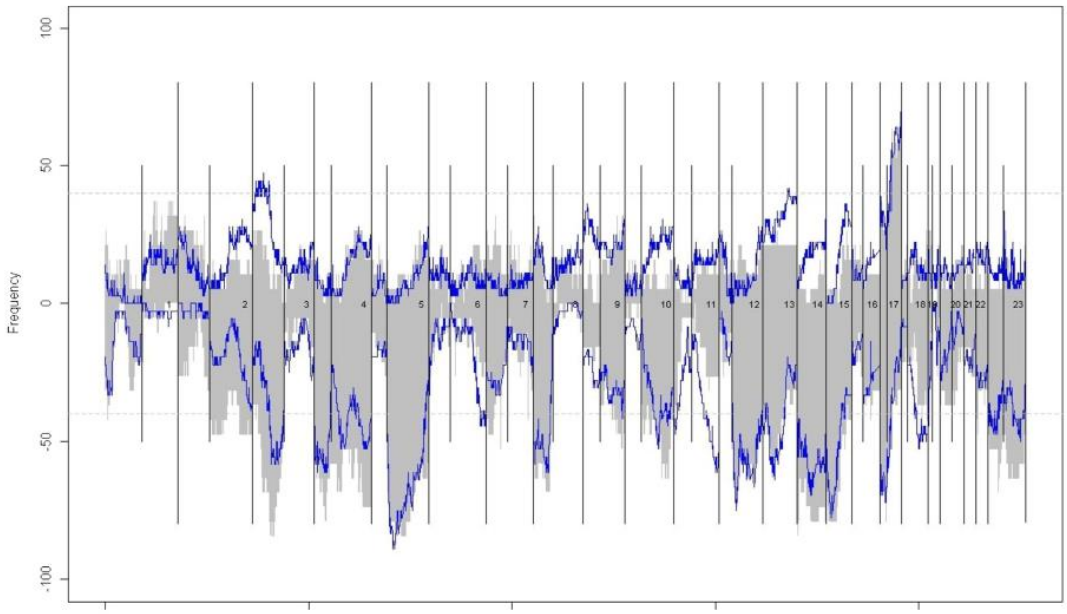


Figure 4 Loss of heterozygosity (LOH) in medullary breast carcinoma (MBC) and non-MBC basal-like carcinoma (BLC). The higher vertical lines indicate chromosome boundaries, the smaller vertical lines indicate centromere boundaries, and horizontal hemidashed lines indicate the 40% frequency threshold. Positive and negative frequencies correspond with frequencies of positive (with gains and copy-neutral) and negative (with losses) LOH.

(Supplemental Table S8). These fusions were not recurrent among the 19 MBCs.

MBC Belongs to the IM Subgroup from the Original TNBC Classification and to the BL1 Subgroup in the Four Classes Refined Classification

To investigate whether MBC belongs to one of the previously defined transcriptomic TNBC subgroups, the MBC transcriptomic profile was submitted to the online tool developed by Lehmann et al.[18] This analysis found that 15 of 18 MBCs belonged to the IM group versus only 3 of 28 non-MBC BLCs by using the original TNBC classification of six classes (Figures 1 and 3). In the refined TNBCtype-4 classification, 11 of 18 MBCs belonged to the BL1 group versus 6 of 28 non-MBC BLCs ($P < 0.05$) (Figures 1 and 3). Interestingly, the three non-MBC BLCs classified as IM in the TNBCtype classification (and as one BL1 and two BL2 in the TNBCtype-4 classification) belonged to the MBC-enriched transcriptomic cluster (Figure 1). Retrospective review of pathology specimens found that two of these tumors could be morphologically be selected as invasive breast carcinoma NST with medullary features, whereas the third could be classified as invasive breast carcinoma NST, with central acellular zone. The stromal signature defined by Lehmann et al[19] did not segregate MBCs from non-MBC BLCs in our cohort (data not shown). Consequently, this study found that the morphologically selected MBC cases belonged, at the transcriptomic level, to the IM and BL1 subgroups as defined by Lehmann et al.[18,19]

MBCs Is Characterized by Frequent LOH

Notably, copy number analysis performed between MBC and non-MBC BLC found, for the first time to our knowledge, 23 recurrent loss of heterozygosity (LOH) specific to MBC and five recurrent LOH specific to non-MBC BLC ($P < 0.05$), listed in detail in Table 7 and plotted as a function of genome location in Figure 4. Recurrent gains and losses are listed in the Supplemental Table S9 and recurrent focal amplicons in Supplemental Table S10. Finally, 11 of 16 MBCs (69%) and 21 of 34 non-MBC BLCs (62%) had high LST status, reflecting *BRCA1/2* inactivation (no significant difference) (Supplemental Table S11).

MBC Has a Better OS and DFS than Non-MBC BLC

All patients in the extended cohort of 32 MBCs and 44 non-MBC BLCs were women between the ages of 27 and 89 years, with a mean age of 53 years. Median clinical follow-up was 101 months (range, 14 to 237 months). Four patients (12.5%) in the MBC group and 13 patients (29.5%) in the non-MBC BLC group died during follow-up. On univariate analysis, MBC, TIL-rich tumors, and N0 cases had a better OS and DFS. On multivariate analysis, MBC subtype and TILs were correlated with a better OS and DFS ($P = 0.02$ for both criteria).

Discussion

This study found that MBC has a specific transcriptomic pattern compared with non-MBC BLC and overexpresses more frequently the long isoform of the *BCLG* gene than other subtypes of breast cancer. *BCLG* overexpression was confirmed at the protein level to be significantly associated with MBC. Moreover, this study found that MBC belonged to the IM subgroup in the six-class system of Lehmann et al[18] and to the BL1 subgroup in the four-class system,[19] respectively. Finally, this study refined the genomic landscape of MBC, harboring more LOH than non-MBC BLC.

The gene expression analysis performed in our series emphasized as expected up-regulation of inflammatory and immune responses pathways in MBC, as previously described in the pioneer study by Bertucci et al.[2] In addition to that, MBC preferentially belonged to the IM TNBC subtype according to the six groups originally defined by Lehmann et al.[18] The latter previously reported that immune signaling genes overexpressed in the IM subtype overlapped with the MBC gene signature as established by Bertucci et al,[2] suggesting that IM could be at least partially represented by MBC.[18] Bareche et al[21] reviewed *a posteriori* the morphology from the TCGA and METABRIC data sets and found an enrichment (15%) of MBC in the IM subgroup. Teschendorff et al[22] found that MBC belonged to an ER-negative gene expression cluster characterized by overrepresentation of genes involved in cell cycle/proliferation and immune response. The favorable outcome of the IM subtype is likely linked to overexpression of immune pathways. Moreover, it is now broadly accepted that TILs are usually associated with a better prognosis in ER-negative cancer.[23] The better outcome of MBC was indeed related to a high TILs infiltrate when compared with non-MBC BLC in our series. Our study found that MBC, selected morphologically, clearly belong to the IM transcriptomic category. Interestingly, it has been suggested that IM may specifically respond to poly (ADP-ribose) polymerase inhibitors and cisplatin-based therapy.[24] Vinayak et al[25] found that the presence of both stromal and intratumoral TILs specifically promoted the response to platinum- and poly (ADP-ribose) polymerase inhibitor–based neoadjuvant chemotherapy in IM-TNBC (PrECOG 0105).

Using the refined classification by Lehman et al[19] based on only four groups, MBC predominantly pertained to the BL1 subgroup, which is characterized by aberrant DNA signaling and repair functions.[18] This finding is in line with the well-documented high prevalence of BRCAness feature in MBC, confirmed in this study through the LST status.

Taken together, our MBC transcriptomic data refined the position of this entity in the taxonomy of TNBC and confirmed the importance of considering both the involvement of T-cell regulation and DNA repair deregulation pathways for further therapeutic approaches to this tumor type. In addition, BRCAness traits in MBC support that, from a biological point of view, MBC is a good candidate to respond to homologous recombination deficiency targeting drugs and immunotherapy. However, given the intrinsic good prognosis of typical forms of MBC, it may be of interest to integrate morphologic and biological specification of the tumor in the clinical management of MBC. It may still be hypothesized that in advanced disease MBC may benefit from the same therapeutic strategies as those developed in TNBC in general.

Overexpression of the long isoform of *BCLG* was found to be one of the major molecular characteristics of epithelial cells in most MBCs but in less than half of the non-MBC BLCs. Interestingly, *BCLG* overexpression was not determined by gene amplification. It is well known that gene expression is associated with copy number changes in breast cancer only in a limited proportion of cases.[26] Of note, this study described a high prevalence of LOH in MBC compared with non-MBC BLC; however, *BCLG* was not included in a recurrent LOH region in MBC. Thus, it can be hypothesized that epigenetic dysregulation may be implicated. Indeed, DNA methylation alterations have been described as associated with *BCLG* variations of expression in different pathological contexts.[27,28] DNA hypomethylation, associated with gene overexpression, is largely prevalent in cancer, especially in TNBC of good prognosis.[29–31] Thus, DNA methylation alterations may participate, at least partially, to *BCLG* overexpression in MBC. It has been previously reported that *BCLG* is down-regulated in breast cancer, regardless of histological or molecular subtype.[32] This new observation emphasizes the significant association of *BCLG* overexpression and MBC. *BCLG* has been extensively studied in terms of its role as a proapoptotic gene, although antiapoptotic functions of *BCLG* have been also reported in cancer cell lines.[33] Notably, *BCLG* mediates apoptosis under the control of several factors, such as *FAU* and maternal embryonic leucine zipper kinase (MELK), in breast cancer cell lines.[34] Lin et al[34] found that the MELK oncoprotein specifically inhibited BCLG-L–induced apoptosis in breast cancer cell lines. Interestingly, Nakamura et al[35] found that the ubiquitin-like protein monoclonal nonspecific suppressor factor β–BCLG complex enhanced apoptosis and inhibited IL-4 production via ERK phosphorylation in T cells. Moreover, transfection studies found that BCLG-L enhanced apoptosis in $CD4^+$ T cells from healthy subjects.[36] Altogether, these findings suggest that BCLG-L may have a proapoptotic role in breast cancers rich in TILs, such as MBC or TILs-rich non-MBC BLC. Furthermore, whether *BCLG* plays a role in increasing the TILs density remains to be determined. Lastly, *BCLG* is a target gene for *TP53*, known to be mutated in up to 100% of MBC.[4] It would be particularly useful to determine whether *BCLG* overexpression in MBC is a compensatory mechanism for *TP53* dysfunction.

In conclusion, this study provides an in-depth molecular portrait of MBC. MBC overexpresses *BCLG* more frequently than non-MBC BLC. Non-MBC BLC

overexpressing *BCLG* is TILs enriched. These data suggest that MBC may represent the end of a spectrum of BLC with enrichment in immune infiltrate together with an increased *BCLG* overexpression. Furthermore, our study found that MBC belongs not only to the IM subgroup from the six categories defined by Lehmann et al[18] but also to the BL1 subgroup from the four group classification of Lehmann et al[19] based on epithelial carcinomatous cells characteristics only. This finding led us to conclude that immune pathways activation together with DNA repair alterations importantly define MBC in regard of non-MBC BLC. TILs activation mechanisms underlying MBC biology should be considered as the main areas of investigation in this specific entity to more clearly understand its better prognosis and propose more appropriate therapeutic approaches.

Supplemental Data

Supplemental material for this article can be found at *https://doi.org/10.1016/j.ajpath.2018.06.021*.

References

1. Pedersen L, Zedeler K, Holck S, Mouridsen HT: Medullary carcinoma of the breast. Prevalence and prognostic importance of classical risk factors in breast cancer. Eur J Cancer 1995, 31A:2289–2295
2. Bertucci F, Finetti P, Cervera N, Charafe-Jauffret E, Mamessier E, Adélaïde J, Debono S, Houvenaeghel G, Maraninchi D, Viens P, Charpin C, Jacquemier J, Birnbaum D: Gene expression profiling shows medullary breast cancer is a subgroup of basal breast cancers. Cancer Res 2006, 66:4636–4644
3. Vincent-Salomon A, Gruel N, Lucchesi C, MacGrogan G, Dendale R, Sigal-Zafrani B, Longy M, Raynal V, Pierron G, de Mascarel I, Taris C, Stoppa-Lyonnet D, Pierga JY, Salmon R, Sastre-Garau X, Fourquet A, Delattre O, de Cremoux P, Aurias A: Identification of typical medullary breast carcinoma as a genomic sub-group of basal-like carcinomas, a heterogeneous new molecular entity. Breast Cancer Res 2007, 9:R24
4. de Cremoux P, Salomon AV, Liva S, Dendale R, Bouchind'homme B, Martin E, Sastre-Garau X, Magdalenat H, Fourquet A, Soussi T: p53 mutation as a genetic trait of typical medullary breast carcinoma. J Natl Cancer Inst 1999, 91:641–643
5. Huober J, Gelber S, Goldhirsch A, Coates AS, Viale G, Öhlschlegel C, Price KN, Gelber RD, Regan MM, Thürlimann B: Prognosis of medullary breast cancer: analysis of 13 International Breast Cancer Study Group (IBCSG) trials. Ann Oncol 2012, 23:2843–2851
6. Ridolfi RL, Rosen PP, Port A, Kinne D, Miké V: Medullary carcinoma of the breast: a clinicopathologic study with 10 year follow-up. Cancer 1977, 40:1365–1385
7. Salgado R, Denkert C, Demaria S, Sirtaine N, Klauschen F, Pruneri G, Wienert S, Van den Eynden G, Baehner FL, Penault-Llorca F, Perez EA, Thompson EA, Symmans WF, Richardson AL, Brock J, Criscitiello C, Bailey H, Ignitiadis M, Floris G, Sparano J, Kos Z, Nielsen T, Rimm DL, Allison KH, Reis-Filho JS, Loibl S, Sotiriou C, Viale G, Badve S, Adams S, Willard-Gallo K, Loi S: The evaluation of tumor-infiltrating lymphocytes (TILs) in breast cancer: recommendation by an International TILs Working Group 2014. Ann Oncol 2015, 26:259–271
8. Chomczynski P, Sacchi N: Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. Anal Biochem 1987, 162:156–159
9. Mi H, Huang X, Muruganujan A, Tang H, Mills C, Kang D, Thomas PD: PANTHER version 11: expanded annotation data from Gene Ontology and Reactome pathways, and data analysis tool enhancements. Nucleic Acids Res 2017, 45:D183–D189
10. Bieche I, Onody P, Laurendeau I, Olivi M, Vidaud D, Lidereau R, Vidaud M: Real-time reverse transcription-PCR assay for future management of ERBB2-based clinical applications. Clin Chem 1999, 45:1148–1156
11. Bieche I, Parfait B, Le Doussal V, Olivi M, Rio MC, Lidereau R, Vidaud M: Identification of CGA as a novel estrogen receptor-responsive gene in breast cancer: an outstanding candidate marker to predict the response to endocrine therapy. Cancer Res 2001, 61:1652–1658
12. Popova T, Manié E, Stoppa-Lyonnet D, Rigaill G, Barillot E, Stern MH: Genome Alteration Print (GAP): a tool to visualize and mine complex cancer genomic profiles obtained by SNP arrays. Genome Biol 2009, 10:R128
13. Popova T, Manié E, Rieunier G, Caux-Moncoutier V, Tirapo C, Dubois T, Delattre O, Sigal-Zafrani B, Bollet M, Longy M, Houdayer C, Sastre-Garau X, Vincent-Salomon A, Stoppa-Lyonnet D, Stern MH: Ploidy and large-scale genomic instability consistently identify basal-like breast carcinomas with BRCA1/2 inactivation. Cancer Res 2012, 72:5454–5462
14. Manié E, Vincent-Salomon A, Lehmann-Che J, Pierron G, Turpin E, Warcoin M, Gruel N, Lebigot I, Sastre-Garau X, Lidereau R, Remenieras A, Feunteun J, Delattre O, de Thé H, Stoppa-Lyonnet D, Stern MH: High frequency of TP53 mutation in BRCA1 and sporadic basal-LIKE carcinomas but not in BRCA1 luminal breast tumors. Cancer Res 2009, 69:663–671
15. Kim D, Salzberg SL: TopHat-Fusion: an algorithm for discovery of novel fusion transcripts. Genome Biol 2011, 12:R72
16. McPherson A, Hormozdiari F, Zayed A, Giuliany R, Ha G, Sun MG, Griffith M, Heravi Moussavi A, Senz J, Melnyk N, Pacheco M, Marra MA, Hirst M, Nielsen TO, Sahinalp SC, Huntsman D, Shah SP: deFuse: an algorithm for gene fusion discovery in tumor RNA-Seq data. PLoS Comput Biol 2011, 7:e1001138
17. Trapnell C, Pachter L, Salzberg SL: TopHat: discovering splice junctions with RNA-Seq. Bioinformatics 2009, 25:1105–1111
18. Lehmann BD, Bauer JA, Chen X, Sanders ME, Chakravarthy AB, Shyr Y, Pietenpol JA: Identification of human triple-negative breast cancer subtypes and preclinical models for selection of targeted therapies. J Clin Invest 2011, 121:2750–2767
19. Lehmann BD, Jovanović B, Chen X, Estrada MV, Johnson KN, Shyr Y, Moses HL, Sanders ME, Pietenpol JA: Refinement of triple-negative breast cancer molecular subtypes: implications for neoadjuvant chemotherapy selection. PLoS One 2016, 11:e0157368
20. Jiang G, Zhang S, Yazdanparast A, Li M, Pawar AV, Liu W, Inavolu SM, Cheng L: Comprehensive comparison of molecular portraits between cell lines and tumors in breast cancer. BMC Genomics 2016, 17 Suppl 7:525
21. Bareche Y, Venet D, Ignatiadis M, Aftimos P, Piccart M, Rothe F, Sotiriou C: Unravelling triple-negative breast cancer molecular heterogeneity using an integrative multiomic analysis. Ann Oncol 2018, 29:895–902
22. Teschendorff AE, Miremadi A, Pinder SE, Ellis IO, Caldas C: An immune response gene expression module identifies a good prognosis subtype in estrogen receptor negative breast cancer. Genome Biol 2007, 8:R157
23. Ibrahim EM, Al-Foheidi ME, Al-Mansour MM, Kazkaz GA: The prognostic value of tumor-infiltrating lymphocytes in triple-negative breast cancer: a meta-analysis. Breast Cancer Res Treat 2014, 148:467–476
24. Abramson VG, Lehmann BD, Ballinger TJ, Pietenpol JA: Subtyping of triple-negative breast cancer: implications for therapy. Cancer 2015, 121:8–16
25. Vinayak S, Gray RJ, Adams S, Jensen KC, Manola J, Afghahi A, Goldstein LJ, Ford JM, Badve SS, Telli ML: Association of increased

tumor-infiltrating lymphocytes (TILs) with immunomodulatory (IM) triple-negative breast cancer (TNBC) subtype and response to neo-adjuvant platinum-based therapy in PrECOG0105. J Clin Oncol 2014, 32:15_suppl, 1000-1000

26. Chin K, DeVries S, Fridlyand J, Spellman PT, Roydasgupta R, Kuo WL, Lapuk A, Neve RM, Qian Z, Ryder T, Chen F, Feiler H, Tokuyasu T, Kingsley C, Dairkee S, Meng Z, Chew K, Pinkel D, Jain A, Ljung BM, Esserman L, Albertson DG, Waldman FM, Gray JM: Genomic and transcriptional aberrations linked to breast cancer pathophysiologies. Cancer Cell 2006, 10:529–541
27. Balasubramanian D, Akhtar-Zaidi B, Song L, Bartels CF, Veigl M, Beard L, Myeroff L, Guda K, Lutterbaugh J, Willis J, Crawford GE, Markowitz SD, Scacheri PC: H3K4me3 inversely correlates with DNA methylation at a large class of non-CpG-island-containing start sites. Genome Med 2012, 4:47
28. Coit P, Jeffries M, Altorok N, Dozmorov MG, Koelsch KA, Wren JD, Merrill JT, McCuneWJ, Sawalha AH: Genome-wide DNA methylation study suggests epigenetic accessibility and transcriptional poising of interferon-regulated genes in naïve CD4+ T cells from lupus patients. J Autoimmun 2013, 47:78–84
29. Ting AH, McGarvey KM, Baylin SB: The cancer epigenome – components and functional correlates. Genes Dev 2006, 20: 3215–3231
30. Jackson K, Yu MC, Arakawa K, Fiala E, Youri B, Fiegl H, Müller-Holzner E, Wolschwendter M, Ehrlich M: DNA hypomethylation is prevalent even in low-grade breast cancers. Cancer Biol Ther 2004, 3: 1225–1231
31. Stirzaker C, Zotenko E, Song JZ, Qu W, Nair SN, Locke WJ, Stone A, Armstrong NJ, Robinson MD, Dobrovic A, Avery-Kiejda KA, Peters KM, French JD, Stein S, Korbie DJ, Trau M, Forbes JF, Scott RJ, Brown MA, Francis GD, Clark SJ: Methylome sequencing in triple-negative breast cancer reveals distinct methylation clusters with prognostic value. Nat Commun 2015, 6:5899
32. Pickard MR, Green AR, Ellis IO, Caldas C, Hedge VL, Mourtada-Maarabouni M, Williams GT: Dysregulated expression of Fau and MELK is associated with poor prognosis in breast cancer. Breast Cancer Res 2009, 11:R60
33. Pickard MR, Mourtada-Maarabouni M, Williams GT: Candidate tumour suppressor Fau regulates apoptosis in human cells: an essential role for Bcl-G. Biochim Biophys Acta 2011, 1812:1146–1153
34. Lin M-L, Park J-H, Nishidate T, Nakamura Y, Katagiri T: Involvement of maternal embryonic leucine zipper kinase (MELK) in mammary carcinogenesis through interaction with Bcl-G, a pro-apoptotic member of the Bcl-2 family. Breast Cancer Res 2007, 9:R17
35. Nakamura M, Nakagawa M, Watanabe J: Ubiquitin-like protein MNSFβ negatively regulates T cell function and survival. Immunol Invest 2015, 44:1–12
36. Luo N, Wu Y, Chen Y, Yang Z, Guo S, Fei L, Zhou D, Yang C, Wu S, Ni B, Hao F, Wu Y: Upregulated BclG(L) expression enhances apoptosis of peripheral blood CD4+ T lymphocytes in patients with systemic lupus erythematosus. Clin Immunol 2009, 132:349–361

Les cellules tumorales engagent souvent des mécanismes existant à l'état physiologique. Les techniques de séquençage à haut débit employées depuis une dizaine d'années sur de vastes cohortes ont permis de mettre à jour l'implication dans la biologie des tumeurs de phénomènes impliqués dans le contrôle de l'expression des gènes. Ainsi, les modificateurs de la chromatine constituent une famille de gènes fréquemment mutés dans les cancers. Le complexe de répression transcriptionnelle Polycomb PRC2 catalyse la méthylation de H3K27, marque associée aux gènes inactifs. PRC2 est l'objet de diverses altérations dans les cancers dont les conséquences sur l'activation des gènes et la croissance tumorale varient selon le type tumoral et la nature de l'altération. Certaines de ces altérations résultent en une abrogation partielle ou totale de la déposition de H3K27me3, alors que d'autres, nommément la mutation Y646 de la sous-unité catalytique de PRC2 EZH2 rencontrée dans environ 25% des lymphomes folliculaires, se caractérise par une augmentation et une redistribution de H3K27me3. PRC2 peut donc être alternativement considéré comme suppresseur de tumeur ou oncogène selon le contexte. Cette thèse a pour objectif de clarifier certains des mécanismes qui sous-tendent ces altérations au travers de deux approches.

1/ Modélisation des principales altérations de PRC2 dans les cancers au sein d'un modèle cellulaire isogénique unique : une approche mécanistique

Nous avons utilisé une lignée de fibroblastes embryonnaires murin immortalisés (Mef) porteurs d'une délétion conditionnelle d'*Ezh2* et récapitulant les génotypes *Eed* KO, H3.3K27M, *Ezh2*$^{Y641/WT}$, *Ezh2*$^{Y641/-}$ +/- *Ezh1* KO. Nous confirmons que : la perte de fonction de PRC2 (totale ou partielle) conduit à une diminution de H3K27me3, les conséquences transcriptionnelles des pertes de fonction partielles de PRC2 sont une réactivation de cibles de PRC2, sans génération de nouvelles cibles liées à H3.3K27M notamment. A l'inverse, *Ezh2*^{Y641} induit une augmentation de H3K27me3, indépendamment de *Ezh2*^{WT} et de *Ezh1*. Nous montrons une participation des cofacteurs de PRC2 dans l'activité enzymatique intrinsèque de *Ezh2*^{Y641}. Nous confirmons la redistribution de H3K27me2/3 dans les cellules *Ezh2*$^{Y641/WT}$ et la restauration d'un profil « WT-like » dans les cellules *Ezh2*$^{Y641/-}$ traitées par inhibiteur d'Ezh2. Les pertes de fonction de PRC2 sont également associées à une augmentation globale de H3K27ac avec diminution relative au niveau des pics, alors que *Ezh2*^{Y641} induit une augmentation globale de H3K27ac à l'échelle de l'ensemble du génome. La réponse transcriptionnelle aux inhibiteurs d'Ezh2 montre un nombre de gènes dé-réprimés similaire mais de nature différente dans les cellules *Ezh2*$^{Y641/WT}$ et *Ezh2*$^{WT/WT}$ à dose égale, sans lien évident avec le niveau de H3K27me3. Enfin, *Ezh2*$^{Y641/WT}$ montre de façon inattendue une réexpression des gènes impliqués dans la présentation antigénique dans les Mef après traitement, de façon similaire à certaines observations faites dans des modèles de lymphomes B centro-germinatifs.

2/ Constitution d'une cohorte longitudinale de lymphomes folliculaires *EZH2* WT et mutés : une approche translationnelle

Le séquençage de l'exon 16 et 18 d'*EZH2* à partir de l'ADN de 160 patients traités pour un lymphome folliculaire entre 1988 et 2017 a permis d'identifier 18.8% de cas mutés. Sur une cohorte restreinte de 32 patients (21 mutés et 11 WT) avec un suivi clinique médian de 11.5 ans, nous avons réalisé, sur au moins un temps biopsique congelé, une étude intégrative associant séquençage couplé et CNV d'environ 600 gènes impliqués dans le cancer (panel DRAGON), RNAseq et ChIPseq H3K27me3. Pour 13 cas, ChIPseq (sur tissu congelé ou FFPE) et RNAseq ont été réalisés de façon séquentielle sur au moins 2 biopsies/patient. Nous confirmons la fréquence élevée des mutations de *KMT2D* et *CREBBP* sans co-occurrence significative avec les mutations d'*EZH2*. La signature transcriptionnelle des cas mutés montre autant de gènes activés que de gènes réprimés, avec un enrichissement dans les voies de la communication avec le microenvironnement pour les gènes réprimés par EZH2^{Y646}. La mutation induit une redistribution de H3K27me3 similaire à celle observée dans les Mef, avec un enrichissement sur les corps des gènes et une diminution sur les régions promotrices. L'enrichissement sur les corps des gènes est plus marqué pour les gènes réprimés par la mutation, les gènes activés ne présentent pas de profile H3K27me3 différents selon le génotype. Enfin, les enhancers spécifiques aux cellules B sont enrichis en H3K27me3 dans les cas mutés, en lien avec une répression transcriptionnelle différentielle.

université PARIS-SACLAY | ÉCOLE DOCTORALE Cancérologie : biologie - médecine - santé (CBMS)

Titre : Modélisation des altérations de PRC2 dans les cancers, vers une meilleure compréhension du rôle de la mutation d'*EZH2* dans les lymphomes folliculaires
Mots clés : Epigénétique, PRC2, Cancer, EZH2, Lymphome folliculaire

Résumé : Le séquençage à haut débit a permis d'identifier dans les cancers le rôle des mécanismes généraux de contrôle de l'expression des gènes dont les modificateurs de la chromatine. Le complexe de répression transcriptionnelle Polycomb PRC2 catalyse la méthylation de H3K27, une marque associée aux gènes inactifs. PRC2 est l'objet de diverses altérations dans les cancers dont les conséquences sur l'activation des gènes varient selon le type tumoral et la nature de l'altération. Certaines de ces altérations résultent en une abrogation de H3K27me3, alors que la mutation Y646 de la sous-unité catalytique de PRC2, EZH2, qui est rencontrée dans 25% des lymphomes folliculaires (LF), induit une augmentation/redistribution de H3K27me3. PRC2 peut donc être considéré comme suppresseur de tumeur ou comme oncogène selon le contexte. Cette thèse a eu pour objectif de clarifier les conséquences mécanistiques de ces altérations au travers de deux approches complémentaires.

1/ Approche mécanistique : Modélisation des altérations de PRC2 dans les cancers dans un modèle cellulaire isogénique : Nous avons utilisé une lignée de fibroblastes embryonnaires murins immortalisés porteurs d'une délétion conditionnelle d'*Ezh2* et récapitulant les génotypes *Eed*-KO, H3.3K27M, *Ezh2*$^{Y641F/WT}$ et *Ezh2* $^{Y641F/-}$ +/- *Ezh1*-KO. La perte de fonction de PRC2 conduit à une diminution d'H3K27me3 et la réactivation de cibles de PRC2. En particulier, nous n'observons pas la génération de nouvelles cibles, ni le gain de cette marque en présence de H3.3K27M comme cela avait été décrit précédemment. De façon intéressante, les pertes de fonction de PRC2 sont associées à une augmentation globale de H3K27ac avec toutefois une diminution relative au niveau des pics. A l'inverse, la mutation gain de fonction *Ezh2*^{Y641F} induit une augmentation de H3K27me3 et ne nécessite pas la présence d'un allèle WT ou de *Ezh1* pour assurer la monométhylation. *Ezh2*^{Y641F} entraîne une redistribution de H3K27me3 dans les cellules *Ezh2*$^{Y641F/WT}$ avec un aplatissement des pics mais une propagation globale de la déposition. Cette altération des propriétés de PRC2 est due à un profond changement de la manière dont PRC2-Ezh2Y641F interagit avec la chromatine car l'inhibition partielle de son activité ne restaure pas un profil de position de la marque sauvage. La réponse transcriptionnelle aux inhibiteurs d'EZH2 est quantitativement similaire entre les lignées WT et *Ezh2*$^{Y641F/WT}$ mais la nature des gènes est différente : la lignée *Ezh2*$^{Y641F/WT}$ montre de façon inattendue une réexpression de gènes impliqués dans la présentation antigénique, comme observé dans des modèles de lymphomes B.

2/ Approche translationnelle : Constitution d'une cohorte de lymphomes folliculaires *EZH2* WT et mutés : Le séquençage des exons 16/18 d'*EZH2* à partir de 160 cas de LF a permis d'identifier 18.8% de cas mutés. Sur une cohorte restreinte de 32 patients (21 mutés et 11 WT) avec un suivi clinique médian de 11.5 ans, nous avons réalisé, sur au moins un temps biopsique, une étude intégrative associant séquençage et CNV de 571 gènes impliqués dans le cancer, RNA-Seq et ChIP-Seq H3K27me3. Pour 13 cas, ChIP-Seq et RNA-Seq ont été réalisés de façon séquentielle sur >2 biopsies/patient. Nous confirmons la prévalence des mutations de *KMT2D* et *CREBBP* sans co-occurrence spécifique avec les mutations d'*EZH2*. La signature transcriptionnelle des cas mutés montre autant de gènes activés que réprimés, avec un enrichissement dans les voies du microenvironnement pour les gènes réprimés. La mutation induit une redistribution de H3K27me3 avec un enrichissement sur les corps des gènes et une diminution sur les régions promotrices. L'enrichissement sur les corps des gènes est plus marqué pour les gènes réprimés par *EZH2*^{Y646}, les gènes activés ne présentent pas de profil différent selon le génotype. Enfin, les enhancers spécifiques aux cellules B sont enrichis dans les cas mutés, en lien avec une répression transcriptionnelle différentielle.

Title : Modeling PRC2 alterations in cancer, a window toward a better comprehension of Y646 EZH2 mutant follicular lymphoma
Keywords : Epigenetics, PRC2, Cancer, EZH2, Follicular lymphoma

Abstract : Next generation sequencing unveiled the pivotal role in cancer of general mechanisms involved in the control of gene expression including chromatin modifiers. The Polycomb repressive complex 2 "PRC2" catalyzes the methylation of H3K27, a mark associated with inactive genes. PRC2 is found to be altered in various cancer types and the consequences of these alterations depend on both tumor type and the nature of the alteration. Some of them result in abrogation of the deposition of H3K27me3, whereas other, namely the mutation Y646 of the catalytic sub-unit of PRC2, EZH2, which is present in up to 25% of follicular lymphoma (FL), leads to both increase and redistribution of H3K27me3. Thus, PRC2 can be alternatively considered as tumor suppressor or as oncogene given the context. This PhD thesis aimed at clarifying the mechanistic consequences of these alterations through two complementary approaches

1/ Mechanistic approach: Modeling main alterations of PRC2 in cancers in an isogenic cell line model: We used an immortalized murine embryonic fibroblast cell line harboring a conditional deletion of *Ezh2* and recapitulating *Eed*-KO, H3.3K27M, *Ezh2*$^{Y641F/WT}$ and *Ezh2* $^{Y641F/-}$ +/- *Ezh1*-KO genotypes. Loss-of-function of PRC2 leads to decrease of H3K27me3 and a reactivation of PRC2 target genes. Consistently, we do not observe generation of new target genes nor gain of the mark upon H3.3K27M mutation in contrast to previous report. Interestingly, loss-of-function alterations of PRC2 are associated with a global increase of H3K27ac yet with a relative decrease of the height of the peaks. In contrast, the gain-of-function mutation *Ezh2*^{Y641F} induces an increase of H3K27me3, regardless of the presence of either *Ezh2* WT allele or *Ezh1* to ensure monomethylation. *Ezh2*^{Y641F} leads to a redistribution of H3K27me3 in *Ezh2*$^{Y641F/WT}$ cells with a flattening of the peaks and a global propagation of the deposition. This alteration of PRC2 properties is due to a profound change in the way PRC2-Ezh2Y641 interacts with chromatin since partial reduction of its activity does not fully restore a WT pattern of deposition. Transcriptional response to EZH2 inhibitors is quantitatively similar between WT and *Ezh2*$^{Y641F/WT}$ cells, but the nature of responsive genes is different: *Ezh2*$^{Y641F/WT}$ cells unexpectedly re express genes involved in antigenic presentation, in a similar way as observed in some B cell lymphoma models.

2/ Translational approach: Constitution and follow-up of a cohort of follicular lymphoma cases WT and mutant for EZH2*:* Sequencing of exons 16/18 of *EZH2* from 160 FL cases allowed the identification of 18.8% mutated cases. In a restricted cohort of 32 cases (21 mutant EZH2, 11 WT) with a median follow-up of 11.5 years, we conducted an integrative study including coupled sequencing and CNV analysis of 571 genes involves in cancer, RNA-Seq and H3K27me3 ChIP-Seq. For 13 cases, ChIP-Seq and RNA-Seq have been performed in >2 longitudinal time-point samples/patient. We confirmed the high frequency of *KMT2D* and *CREBBP* mutations without specific co-occurrence with *EZH2* mutations. Transcriptomic signature of *EZH2*^{Y646} cases showed a similar number of *up* and *down* regulated genes, with an enrichment in pathways involved in microenvironmental crosstalk for the repressed genes. *EZH2*^{Y646}induces a redistribution of H3K27me3 along with an enrichment at the gene bodies and a depletion at the promoter regions. Enrichment at the gene bodies is more marked on the downregulated genes in the mutant background, whereas upregulated genes do not show a noticeably different H3K27me3 pattern according to the genotype. Finally, B cell specific enhancers are enriched in H3K27me3 in *EZH2* mutant cases, along with differential transcriptional repression compared with WT cases.

Maison du doctorat de l'Université Paris-Saclay
2ème étage aile ouest, Ecole normale supérieure Paris-Saclay
4 avenue des Sciences,
91190 Gif sur Yvette, France