Université de Montréal

BMI1 mediated heterochromatin compaction represses G-quadruplex formation in Alzheimer's disease

Par

Roy Hanna

Programme de biologie moléculaire, Faculté de médecine,

Université de Montréal

Thèse présentée en vue de l'obtention du grade de doctorat

en biologie moléculaire

Avril, 2020

Résumé

La maladie d'Alzheimer (MA) est la démence la plus importante dans le monde développé. Cette maladie neurodégénérative rend de plus en plus difficile la capacité d'accomplir les tâches quotidiennes de routine, elle peut également faire oublier les mots aux patients, les désorienter dans le temps et l'espace, et à des stades avancés entraîne une perte de mémoire. Malheureusement, la MA est considérée comme le prochain grand défi pour la santé publique de la plupart des pays, le nombre de cas devant doubler au cours des 20 prochaines années en raison du vieillissement de la population. Cette augmentation du nombre de patients s'accompagne d'une augmentation des besoins de financement et de personnel de santé afin de répondre aux demandes et aux besoins de ces patients. La MA peut être divisée en deux entités distinctes: une maladie héréditaire bien définie et bien comprise qui représente jusqu'à 5% de tous les cas de MA appelés maladie d'Alzheimer familiale, et une maladie moins définie appelée maladie d'Alzheimer sporadique. Le facteur de risque le plus défini pour la MA est l'âge, mais récemment, il a été démontré que le cerveau des patients atteints de MA avait un niveau réduit de BMI1 et que la suppression de BMI1 dans les neurones humains ou chez la souris déclenche les caractéristiques de cette maladie.

Alors que BMI1 était connu pour être important dans les stades de développement, nous rapportons ici qu'il est crucial dans les cellules adultes pour maintenir la compaction de la chromatine et l'inhibition de la transcription des séquences répétitives. De plus, ces deux fonctions de BMI1 empêchent l'ADN d'acquérir une conformation G4. Cette conformation peut entraîner une instabilité du génome, une augmentation des dommages à l'ADN et une altération de l'expression des gènes, mais surtout, nous avons montré que dans les neurones corticaux, les structures G4 peuvent influencer l'épissage alternatif de divers gènes, notamment APP. Ces résultats apportent un éclairage nouveau sur l'origine de la maladie et l'importance de BMI1 et de la structure secondaire de l'ADN dans le cadre de la MA.

Mots-clés : BMI1, G-Quadruplex, maladie d'Alzheimer, polycombes, séquences LINE, hétérochromatine, instabilité génomique, G4.

Abstract

Alzheimer's disease is the most prominent dementia in the developed world. This neurodegenerative disease renders the ability to do the routine daily tasks more and more difficult; it can also cause patients to forget words, be disoriented in time and space, leading to a memory loss. Unfortunately, AD is considered the next big challenge for most country's public health, with the number of cases thought to be doubling within the next 20 years due to the aging of the population. This increase in the number of patients comes with an increase in the need for funding and for healthcare personnel to meet the demands and the requirements of these patients. AD is divided into two separate entities: a well-defined and understood hereditary disease that makes up to 5% of all AD cases called familial Alzheimer disease, and a less defined one called sporadic Alzheimer disease. sAD most defined risk factor is age, but recently it was shown that brains of sAD patients had a reduced level of BMI1 and that the knockdown of BMI1 in human neurons or mice triggers the hallmarks of this disease.

While BMI1 was known to be important in the developmental stages, we report here that it is crucial in adult cells to maintain the compaction of the chromatin and the silencing of the repetitive sequences. Furthermore, these two functions of BMI1 prevent the DNA from acquiring a G4 conformation. This conformation can lead to genome instability, increased DNA damage, and altered gene expression. However, most importantly, we showed that in cortical neurons, G4 structures could influence the alternative splicing of various genes, notably APP. These results shed new light on the origin of AD, and the importance of BMI1 and the secondary structure of the DNA in its context.

Keywords: BMI1, G-quadruplex, Alzheimer's disease, polycomb, LINE sequences, heterochromatin, genome instability, G4.

Table of contents

List of tables

List of figures

List of acronyms and abbreviations

3D: Three dimensional

A: Adenine

A.U: Arbitrary unit

Ac: Acetylation

AD: Alzheimer disease

ADP: Adenosine diphosphate

AI: Artificial intelligence

AICD: APP- intracellular domain

ALS: Amyotrophic Lateral Sclerosis

ALU: Arthrobacter luteus restriction endonuclease elements

AMD: Age related macular degeneration

AMP: Adenosine monophosphate

AOR: Antioxidant response

APP: Amyloid precursor protein

ARID: AT-rich interaction domain

ATP: Adenosine triphosphate

ATRX: ATP-dependent helicase ATRX

Av: Antigenic variation

Aβ: Amyloid Beta

BCL2: B-cell lymphoma 2

BLM: Bloom's protein

BMI1: B Lymphoma Mo-MLV insertion region 1 homolog

bp: base pair

BRCA1: Breast cancer type 1 susceptibility protein

BRCA2: Breast cancer type 2 susceptibility protein

BS: Bloom's syndrome

C: Cytosine

CAGE: Cap analysis gene expression

CAP: Cyclic AMP receptor protein

CD133: Cluster of differentiation 133

CDK4/6: Cyclin-dependent kinase 4 and 6 complex

ChIP: Chromatin immuno-precipitation

CIHR: Canadian Institutes of Health Research

cLAD: Constitutive lamina associated domains

CNS: Central nervous system

CSRCS: Comité de Surveillance de la Recherche sur les Cellules Souches

DDR: DNA damage response

DIV: Days in vitro

DKN: Double knockdown

DM2: Myotonic Dystrophy 2 gene

DNA: Deoxyribonucleic acid

DNMT: DNA methyltransferase

DRB: 5,6-Dichloro-1-β-d-ribofuranosylbenzimidazole

E: Glutamic acid

e18.5: Embryonic day 18.5

EOAD: Early onset Alzheimer's disease

ERV: Endogenous retroviruses

ES cells: Embryonic stem cells

ETS1: Protein C-ets-1

FA: Fanconi anaemia

fAD: Familial Alzheimer's disease

FANCJ: Fanconi anemia group J protein

fLAD: Facultative lamina associated domains

FMRP: Fragile X mental retardation protein

FPLC: Fast protein liquid chromatography

FTD: Frontotemporal dementia

G: Guanine

G4: G-quadruplex

GEO: Gene Ontology

GFP: Green fluorescent protein

G-overhang: Guanine rich overhang

GTEx: Genotype-Tissue Expression

GWAS: Genome-Wide Association Studies

Gy: Gray

H: Hydrogen

HCL: Hydrochloric acid

HDAC: Histone deacetylase

HDACi: Histone deacetylase inhibitor

HDFs: Human diploid fibroblasts

hESC: Human embryonic stem cells

HGPS: Hutchinson-Gilford progeria syndrome

HMR: Maisonneuve-Rosemont Hospital

HOXC13: Homeobox C13

HP1: Heterochromatin protein 1

HPA: Human Protein Atlas

HR: Homologous recombination repair

hTERT: Human telomerase reverse transcriptase

HUSH: Human Silencing Hub

IF: Immunofluorescence

IHC: Immuno-histochemistry

IP: Immunoprecipitation

iPSC: Induced pluripotent stem cells

kb: kilo bases

Kda: Kilo daltons

KO: Knock out

LAD: Lamina associated domains

LC-MS: Liquid chromatography–mass spectrometry

LINE: Long interspersed nuclear repeats

LOAD: Late onset Alzheimer's disease

LTR: Long terminal repeats

MA: Maladie d'Alzheimer

MAPT: microtubule-associated protein tau

MDM2: Mouse double minute 2

Me: Methylation

MEME: Multiple Expectation maximization for motif discovery

MIR: Mammalian-wide interspersed repeats

MLL1: Myeloid/lymphoid leukemia protein 1

MnSOD: Manganese superoxide dismutase

MRI: Magnetic resonance imaging

NCBI: National Center for Biotechnology Information

NER: Nucleotide excision repair

NFT: Neurofibrillary tangles

NLS: Nuclear localization signal domain

NSERC: National Science and Engineering Research Council of Canada

O: Oxygen

O.C.T: Tissue-Tek® optimum cutting temperature

ORF: Open reading frame

P: Proline

P30: Day 30 postnatal

PcG: Polycomb-group proteins

PCH: Pericentromeric heterochromatin

PCR: Polymerase chain reaction

PEST: Proline, glutamic acid, serine, and threonine rich domain

PEV: Position-effect variegation

PFA: Paraformaldehyde

PHD: Plant homeodomain

Pifi: Pifithrin

Pol: Polymerase

PQS: Putative quadruplex sequences

PRC1: Polycomb repressive complex 1

PRC2: polycomb repressive complex 2

PSEN1: Presenilin 1

PSEN2: Presenilin 2

Pu: Purine

Py: Pyrimidine

Pyrido: Pyridostatin

QGRS: Quadruplex forming G-rich sequences

qPCR: Quantitative Polymerase chain reaction

Rb: Retinoblastoma

RIDGEs: Regions of increased gene expression

RING : Really Interesting New Gene

R-Loops: DNA:RNA hybrid loop

RNA: Ribonucleic acid

RNAPII: RNA polymerase II

ROS: Reactive oxygen species

RPA: Replication protein A

RRV: Rhesus macaque rhadinovirus

RT: Real time

RYBP: RING1 and YY1-binding protein

S: Serine

sAD: Sporadic Alzheimer disease

SANT: Swi3, Ada2, N-Cor, and TFIIIB

SB: Sodium Butyrate

SCL: Stem cell leukemia gene

SET: Su(var)3-9, Enhancer-of-zeste and Trithorax

Sgs1: Slow growth suppressor 1

shBMI1: Short hairpin targeted against BMI1

shRNA: Short hairpin RNA

shSCR: Short hairpin targeted against a scramble sequence

SINE: Short interspersed nuclear repeats

siRNA: Small interfering RNA

SMNDC1: Survival motor neuron domain containing 1

SPM: Sterile alpha motif

SSM: Slipped-strand mispairing

STR: Short tandem repeats

T: Thymine

TAD: Topological associated domains

Taq: Thermus aquaticus

TFIIH: Transcription factor II Human

T-loop: Telomeric loop

TMEJ: Polymerase theta-mediated end joining

TSA: Trichostatin A

TSS: Transcription start site

Ub: Ubiquitination

UTR: Untranslated region

WHO: World health organization

WRN: Werner helicase

WS: Werner's syndrome

WT: Wild type

XP: Xeroderma pigmentosum

XPB: Xeroderma pigmentosum type B

XPD: Xeroderma pigmentosum type D

ZBP: Z-DNA binding proteins

For my family

Acknowledgments

Foremost, I would like to present my sincere gratitude to my thesis director, Professor Gilbert Bernier. For welcoming me into his lab and ensuring a prosperous and nurturing environment that helped me realize this thesis. I will always be thankful for all the scientific discussions that we had and that helped broaden my horizon and sharpen my knowledge. Furthermore, Dr. Bernier allowed me to discover and to improve my skills in many fields, encouraging me to attend courses and seminars on programming and artificial intelligence that were somehow out of the scope of a traditional molecular biology course, but that benefited me to the fullest. I would like to thank as well my Ph.D. jurors, Dr. Hugo Wurtele, Dr. Julie Lessard, Dr. David Picketts, and Dr. Nicole Francis, for taking the time to read my work and for their valuable inputs.

Alongside I would like to thank my lab members, especially Dr. Anthony Flamier and Andrea Barabino, who have been there for me, supporting me morally and practically with the experiments. The long hours we've spent discussing and analyzing are priceless, not just for this work but also for me personally.

My sincere thanks also go to all my friends in Montreal that have become like a family of mine, your presence in my life these past years is priceless. A web of inter-support that is essential for the success of us all.

Finally, all this work would not have been possible without the full support of my family, my parents, my sisters, and my cousins that trusted in me and supported me unconditionally. I thank you so much for all the sacrifices that you have made, for encouraging me in all of my pursuits and inspiring me to follow my dreams, and for you, I dedicate this work.

1 Chapter 1 – Introduction

In this thesis, we are going to go through a journey from the molecular to the whole organism passing by the microscopic to understand the complexity of Alzheimer's disease better. We are going to start our journey with the molecular structure of the DNA. The next step will touch the macroscopic scale, studying the cell and its aging process. Building on our knowledge from the first step, we will understand the influence of the molecular scale on the cell. Finally, we will introduce Alzheimer's disease, an illness affecting the brain studying it from a macroscopic point of view, but also exploring the influence of the previous two steps on the onset of this disease and its regulation.

1.1 Secondary structure of the DNA

In this section, we are going to introduce the basis of secondary structures of the DNA and their known roles. The secondary structure of the DNA refers to how molecules of DNA are organized along the fiber axis of this polymer.

1.1.1 Complementary helix structure

The DNA is the macromolecule that holds the genetic information for all living organism (1,2). The building blocks of this polymer are four nucleotides (monomers) classified in two groups: Purines with adenine (A) and guanine (G), and the pyrimidines with Cytosine (C) and thymine (T) (3). These nucleotides are organized in two sugar-phosphate chains running in opposite directions (4). The two chains are linked together with hydrogen bonds between the bases of these nucleotides, and hydrogen bonds create the rungs of the "ladder". The nucleotides pairing is specific, that is why we talk about a complementary structure, with adenine paired only with thymine and guanine with cytosine (5). This complementary property was established by Chargaff's law, which states that in each cell, the quantity of adenine is equal to that of thymine, and the amount of guanine is identical to that of cytosine (5). This model does not specify the order of the nucleotides along the chains.

While all the above described the chemical composition of the DNA, in this essay, we are more interested in the secondary structure of this polymer. By secondary structure, we are referring to the unique organization of the molecules in the 3D space. This ladder is twisted along its axis to form a double-stranded helix (6,7). Hence many forms of DNA arise dependent on the helix chemical and biophysical properties. From these forms, three are biologically relevant: A-DNA, B-DNA, and Z-DNA (8).

	A-DNA	B-DNA	Z-DNA
Helix sense	Right-handed	Right-handed	Left-handed
Rotation per bp (twist)	+33°	+36°	-30°
Sequence	Random	Random	Pu/Py alternation
Base pair per turn	11	10	12
Length of a complete helix (pitch)	28 Å	34 Å	45 Å
vertical rise per bp	2.56 Å	3.4 Å	3.7 Å
helical diameter	26 Å	20 Å	18 Å

Table 1 . Physical properties of the DNA double helix

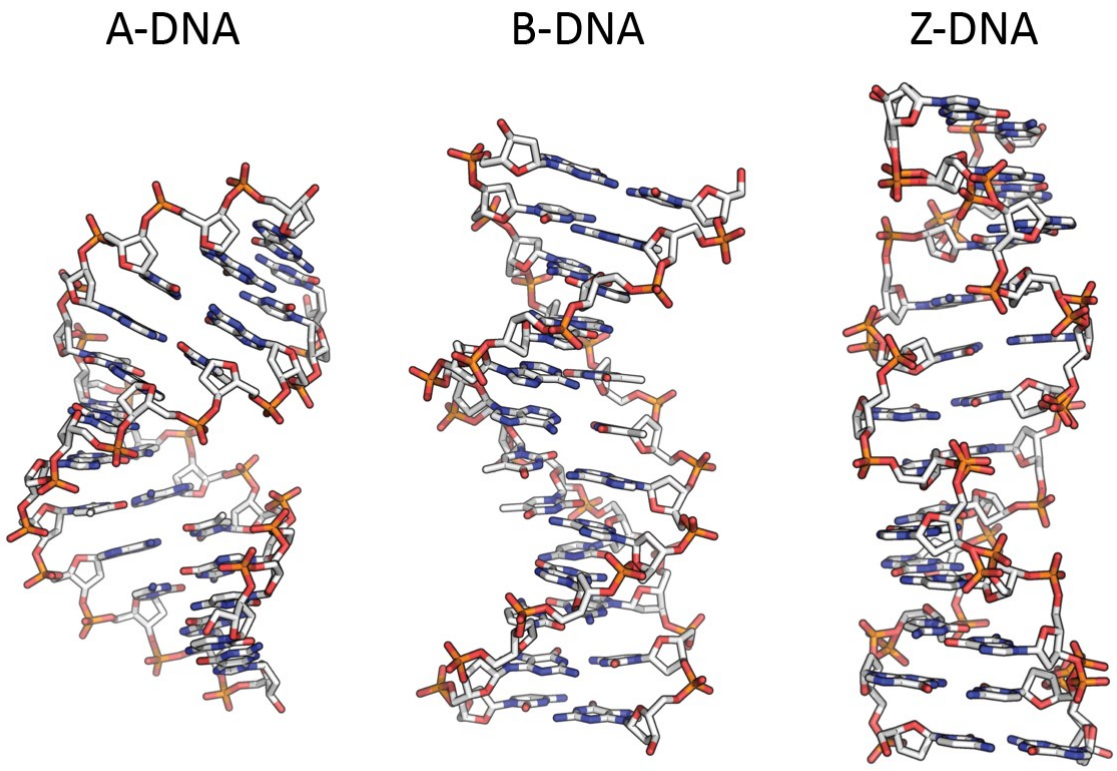


Figure 1. 3D representation of the three forms of DNA (9)

1.1.1.1 B-DNA double helix conformation

Dickerson described B-DNA or the paracrystalline form in 1980 (10). This form represents the vast majority of DNA in living organisms since from an energetic point of view, this form is very stable and neutral (11).

1.1.1.2 A-DNA double helix conformation

Wang described A-DNA or the crystalline form in 1982 (12,13). This form is more compacted then the B-DNA, as we can see in table 1 and fig. 1 (8) with a shorter pitch and a larger diameter. A-DNA was thought to be irrelevant in a biological context because scientists did not have the methods to detect it and quantify it in situ. Whelan, in 2014, provided the first evidence that A-DNA was present in situ in significant quantities (14). This discovery opened the possibility for further research about the role of A-DNA.

1.1.1.2.1 The biological role of A-DNA

Once A-DNA was proven to be relevant in vivo, researchers started to investigate what role it can play. The first hint of its role stemmed from the form of the DNA: a compacted or crystalline form was formed in dehydrated conditions (7,12). This conformation of the A-DNA provides protective properties to the DNA, such as the protection against dissection under dry conditions in the spore formation phase of a bacillus via its interaction with α/β small acid-soluble spore proteins (14–16). Furthermore, A-DNA was proven to also protect against UV radiation (15,17,18). A-DNA can even protect the genome from 5-Bromodeoxyuridine toxicity (19). All these pieces of evidence made the A-DNA structure a potent mediator of genomic stability and protection.

When Mustak et al. probed the brains of suicidal victims of Bipolar Depression, they found an unusually high quantity of A-DNA conformation compared to control patients (20). Their work, even though it does not explain the mechanism behind this association, establishes a clear role of DNA conformation to human disease.

1.1.1.2.2 Protein interactions with A-DNA

Scientists also tried to establish if some proteins interacted specifically with A-DNA. The first to be found is the α/β small acid-soluble spore proteins (14–16) mentioned beforehand. RNA and DNA-RNA helix are known to have the A-form (21,22). For this reason, the reverse transcriptase of HIV interacts with the A-DNA form (23). Cyclic AMP receptor protein (CAP) can bind the motif GTGNxCAC (24), the X in this motif can be a 6 bp or an 8 bp (25). It has been proven that CAP interacts with A-DNA (26,27). That interaction is only possible in the motif with an 8 bp because the A-form shortens the distance between both sides of the motif allowing the protein to interact

(25). Another essential protein that preferentially interacts with A-DNA is the Taq polymerase (Taq Pol) (28,29). The interaction between A-DNA and the DNA pol increases the fidelity of the polymerase replicating the DNA by reducing the sequence-dependent structural alterations of the DNA and thus decreasing the sequence-directed mutations (29–32).

1.1.1.3 Z-DNA double helix conformation

Wang first described Z-DNA conformation in 1979 (33–35). This conformation is more stretched then the previous two with a more extended pitch and a narrower helix diameter, and another big difference is the direction of the twist: A- and B-DNA are right-handed helix while Z-DNA is a left-handed Helix, see table 1 and Fig. 1 (8). Though many doubted the existence of a double helix in the left-handed model because they could not fit it in the original model of DNA, Crick scolded them, referring to their inability to fit the Z-DNA in the model as a "lack of ingenuity on their part" (7). Fortunately, scientists kept searching for antibodies that recognize these structures and developed some that are highly specific (36–38). Using these antibodies, scientists were able to visualize Z-DNA in vivo (39), proving their biological relevance.

1.1.1.3.1 The biological role of Z-DNA

Z-DNA occur in vivo as a transient state (40) in purine pyrimidine alternating regions (41–44). The ephemeral nature of this conformation is a consequence of the Z-DNA being a high energy conformer; this means that Z-DNA requires energy to be formed (11,45) and stabilized (46). Epigenetics changes can provide the necessary conditions for the formation of Z-DNA, specifically the methylation of DNA (47) and histone acetylation (48). Furthermore, many proteins have been discovered that binds Z-DNA; these proteins are called ZBP for Z-DNA binding proteins with the characteristics of having Zα and Zβ domains that have a high affinity to Z-DNA (49–51). Proteins that have the Zα includes ADAR1 (49,50,52), ESL3 (53), E3L (49), Topoisomerase II (54,55). ZBP also plays a role in changing the conformation of the DNA and inducing a B-Z transition (56).

The energy required for this confirmation can also be produced by the passing RNA-pol during transcription that creates a supercoiled DNA with higher energy potential (57,58). Z-DNA is not randomly distributed on the genome but instead concentrated around the transcription start sites (TSS), and these structures were ten times more likely to be near the promoter (5' of the gene)

then to be at the end of the gene (3' of the gene) (59). The first to introduce the model of B-Z transition driven by transcription was Liu in 1987 (58). An elegant experiment in ciliated protozoa reinforced this model. The hypotrichous ciliate *Stylonychia mytilus* has a macronucleus where the transcription occurs during vegetative growth and a micronucleus that is transcriptionally inert (60). When probed with an antibody recognizing Z-DNA, only the macronucleus showed immunoreactivity, establishing the necessity of transcription to the formation of Z-DNA (61). It was not until 1991 that Wittig had the first evidence of the Lui's model in permeabilized mammalian cell nuclei (62). Following this discovery, Z-DNA structures in the promoters of many genes were shown to be proportional to the expression level of these genes, notably: corticotropin hormone-releasing gene (63) and *c-MYC* (64,65). Moreover, a ChIP-seq experiment confirmed the relationship between Z-DNA and transcription (66). It is worth noting that while transcription can induce Z-DNA, in some cases, Z-DNA was found to inhibit transcription by blocking the RNA polymerase and thus may be acting as a negative feedback loop (67,68).

1.1.1.3.2 Relationship between Z-DNA and human diseases

Contrasting the compact structure of the A-DNA that confers it the protective properties of the genome, Z-DNA is stretched and thus theoretically more prone to damage. The susceptibility to DNA damage was proven when it was shown that in human tumors, the chromosomal breakpoints are associated with Z-DNA (69–72) and that this association is because Z-DNA favors recombination (73) because this conformation expose base residues, facilitating any interactions between the two DNA molecules (74). and acts as an anchor for large scale deletions (75). Furthermore, Z-DNA was found near many promoters of tumor genes controlling their expression level like *BCL2*, *SCL* (76), and c-*MYC* (64,65).

Z-DNA is also very immunogenic, which is part of the reason why it was easy to develop specific antibodies against it (36–38). Being highly immunogenic also meant that it was susceptible to induce disease, especially autoimmune disease. Antibodies targeting Z-DNA were found in many autoimmune disorders like lupus erythematosus (36,77), rheumatoid arthritis (78), and Crohn's disease (79). The role of these antibodies is still to be determined as we don't know how they interact with the disease and if they cause it.

Z-DNA was also found to be implicated in many neurodegenerative diseases like Aicardi-Goutières (80), age-related macular degeneration (AMD) (81), Huntington disease and spinocerebellar ataxia (82), Amyotrophic Lateral Sclerosis (ALS) (79) and Alzheimer disease (AD) (83,84).

It is not all bad in the Z-DNA world, this structure was shown to reduce the potential of slipped-strand mispairing (SSM) DNA formation by 62% in the DM2 gene, and thus providing protection against myotonic dystrophy (85). Slipped-strand occurs when we have a large portion of repetitive DNA, mispairing between the complementary strands inducing a "slip" along the helical axis where a part of the genome is not replicated or replicated twice (86). One can argue that the stretch form of the Z-DNA provides more space between the nucleotide, making it harder for a slip to happen.

1.1.2 G-quadruplex

While most of the DNA in the cells is in the double helix form, other conformations may arise, which is the case of G-quadruplex (G4) first described in 1962 by Gellert (87). G4 does not follow the complementary rules of bases, like in the double helix form. G4 is due to the unique ability of guanine to self-assembly by a Hoogsteen hydrogen monovalent bond stabilized by a cation like potassium or sodium in the middle forming a G-quartets (see Fig. 2A) (88). In order to form a proper G4, at least two G-quartets should be stacked on top of each other's (89). This kind of structure can be unimolecular (intramolecular) formed by the same strand of DNA (Fig. 2B), or it can be the result of the assembly of two (bimolecular) (Fig. 2C) or even four (tetramolecular) strands of DNA (Fig. 2D) (90). It is worth noting that while in this essay, we will be talking mostly about DNA G4, RNA can also form these structures (91).

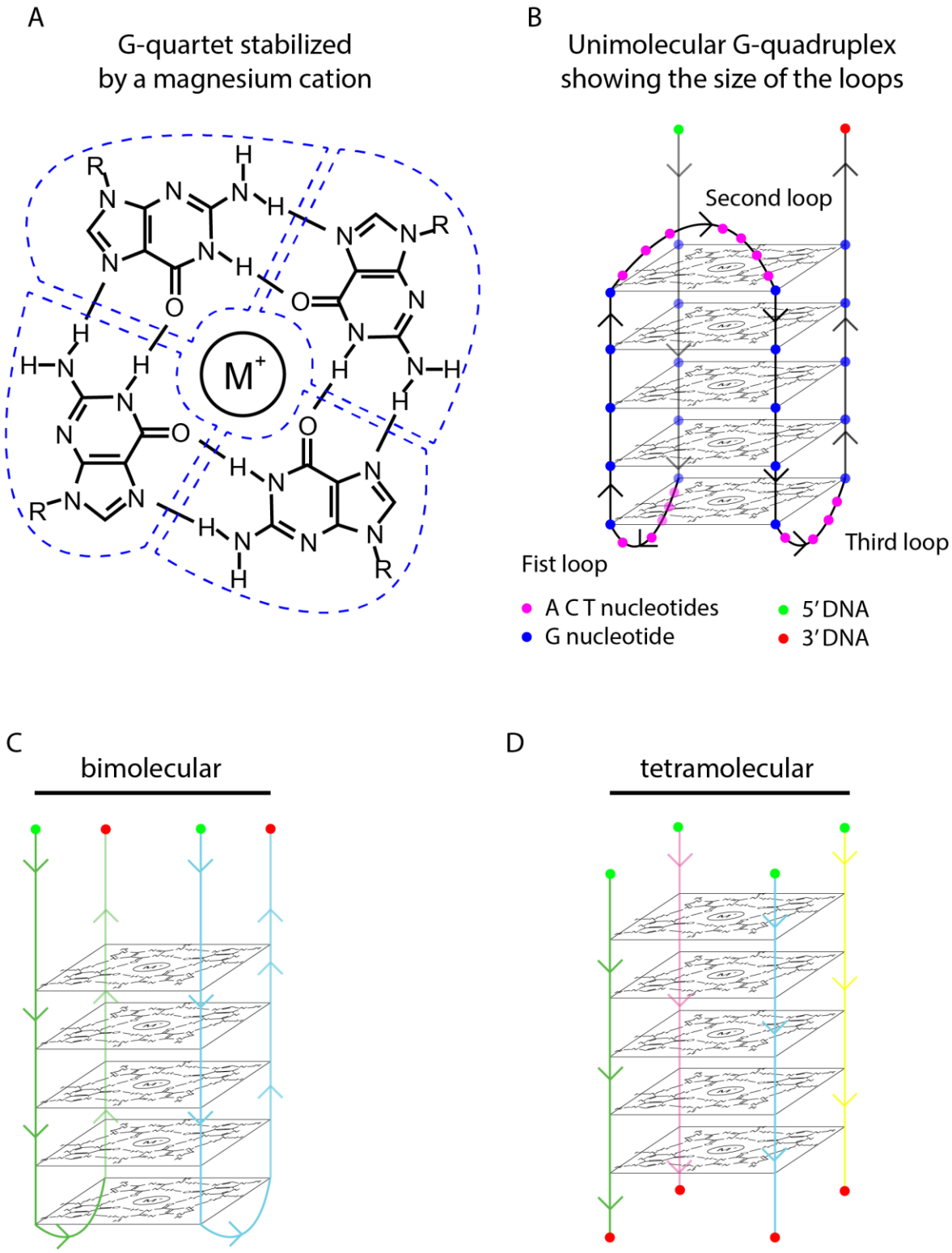


Figure 2. Model of a G4 structure (modified image from (92))

In this study we will be focusing on the unimolecular G4 structures, this category follows a very defined consensus sequence of four stretches of Gs separated by three loops of variable size summarized by $G_m X_n G_m X_n G_m X_n G_m$ where the M value is constant with a minimum of 2 and the N value is variable and can be from 3 to 12 (see Fig. 2B for an example) (93). G4 structures are energetically very stable (90,94); this stability is partially influenced by the length of the loop that was determined for up to 7 bases (95–97). However, this consensus was challenged and then extended to 12 bases because the stability also depends on the cation in the middle of the quartets: small loops are stabilized by potassium, but sodium stabilizes better the long loops (93). The length of the loop does affect not only the stability of the structure but also the topology, that is why we can have G4 structures with an antiparallel (Fig. 3A), parallel (Fig. 3B), or even a mixed topology (Fig.3C)(98).

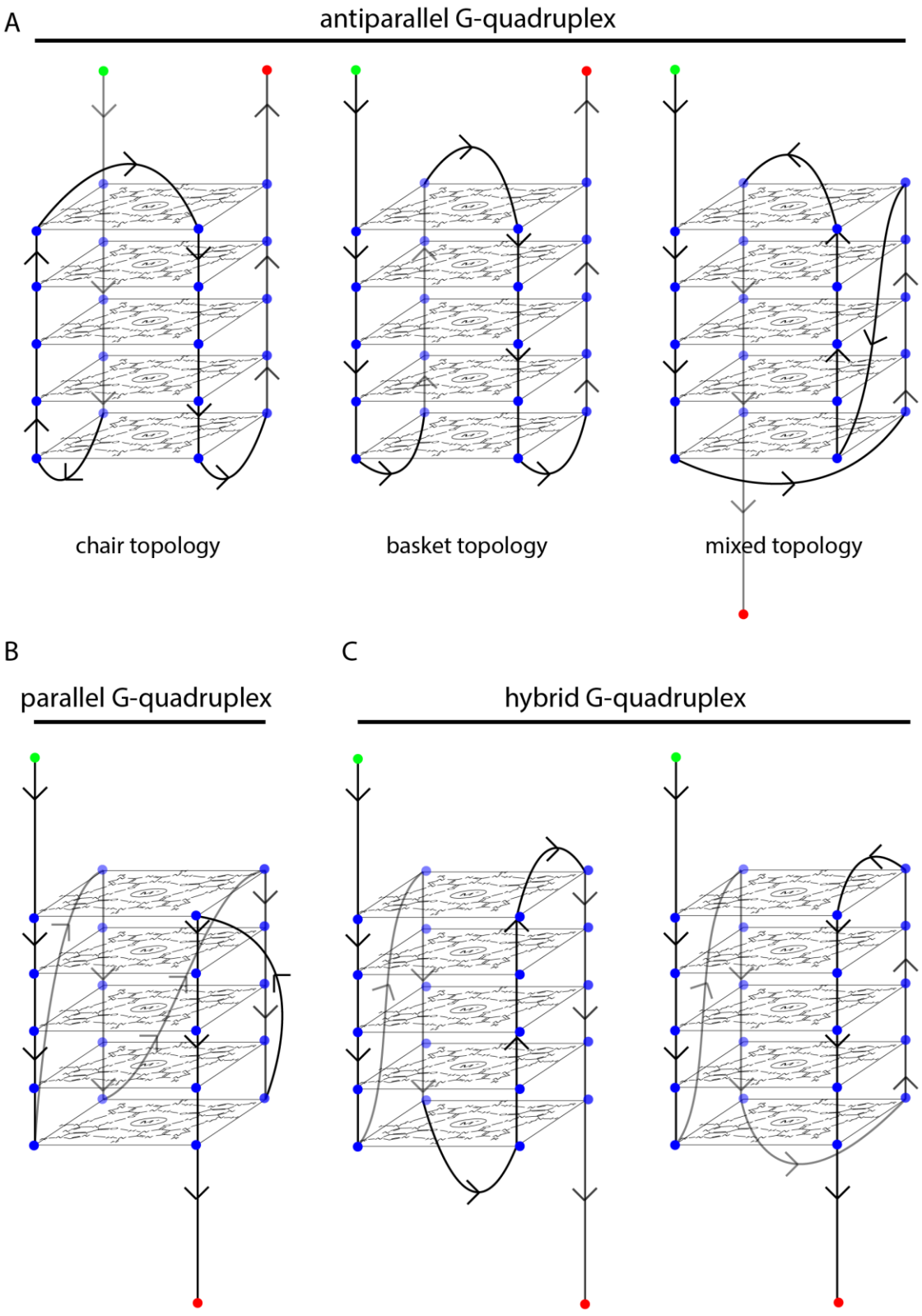


Figure 3. The topology of a unimolecular G4 structure

1.1.2.1 Detection of G4 structures in vivo

1.1.2.1.1 Tools for detecting G4 structures

Since G4 structures are formed with a consensus sequence, many bioinformatics tools have been developed in order to probe the genome and identify the putative G4 sequences in the genome (PQS). PQS are regions on the genome that respect the consensus sequence and can, in theory, form a G4 structure. As an example of this type of software, we can cite the G4hunter (99) or the Quadparser (100). Most recently, artificial intelligence (AI) was trained on a large set of experimental datasets in order to identify the folding capacity of the DNA (101).

Since these algorithms can only identify PQS and not actual structures in the cell, many molecules have been developed to visualize in vivo these structures: IMT (102), and red-NIR (103) are two molecules that can bind nuclear DNA G4 structures. These molecules are not fluorescent except when they interact with the G4 structures, making them valuable tools for probing in vivo G4 structures. Another molecule, the CyT, can selectively bind and label RNA G4 structures (104).

Antibodies that can selectively recognize G4 structures were also developed in order to facilitate the studies in this field. The first antibody was a single-chain variable fragment published in 2008 called hf2 (105), which had a high affinity to G4 structures. Another single-chain antibody clone was detected using the phage display technique in 2013, and this antibody is called BG4 (106). BG4 can recognize both nuclear DNA G4 and cytoplasmic RNA G4 structures (107). In 2014, the 1H6 monoclonal antibody was published with a high affinity to a broad G4 structure conformation and topology: unimolecular or tetramolecular G4 in various topologies (108).

1.1.2.1.2 Localization of G4 structures in the genome

In an early experiment using immune histochemistry experiments (IHC), G4 structures were mistakenly reported to be localized at constitutive heterochromatin (109), but many studies challenged these findings (110). Many publications after, including two ChIP-seq experiments (111,112) showed that G4 structures are present at: 1) constitutive heterochromatin, specifically at telomeres (113–116) to help in capping them (117); 2) origins of replication (118,119); 3) and a significant portion of G4 structures are at the promoter of genes (120,121). In fact, other than the article mentioned above (109), little is known so far about the interplay between chromatin

structure and G4 structures. Most of the work done on G4 structure focused on the proteins that can resolve them as we will see in the next chapter.

1.1.2.2 Modulators of G4 structures in vivo

Since G4 structures have a very stable conformation, their unwinding is an active process with the involvement of specialized proteins. The first helicase to be identified was in *Saccharomyces cerevisiae*: the Sgs1 (122), the homologue of the human BLM protein. Pif1 (123,124) and its human homologous protein PFH1 (125) were also involved in unwinding G4 structures in vivo as well as ATRX (126) and replication protein A (RPA) (127). Moreover, many helicases family were shown to be able to unwind G4 structures like the RECQ family helicases (128,129) which include WRN (130), FANCJ (131–134), and Bloom (BLM) (135), another family is the helicases in the TFIIH complex: XPD and XPB (136). While helicases are a very efficient and faithful way of resolving G4 structures, if such a structure is persistent DNA damage response (DDR) machinery can be involved in excising the G4 structure and employing either homologous recombination repair (HR) (137) or Polymerase theta-mediated end joining (TMEJ) (138).

1.1.2.3 The biological relevance of the G4 structures

Once the existence of the G4 structures in vivo was confirmed, and the localization of these structures was determined and found to be very specific and peculiar, the next question was about the role of these structures at these locations. The function of G4 was very diverse and dependent on the location of it, for example, the G4 present at origins of replications were shown to be required to initiate DNA replication at these sites (139,140). The G4 present at the promoter of these genes was shown to be able to regulate their expression (141,142). For example, in the transcription activation site of the human c-*KIT* oncogene, a G4 structure has been identified and conserved across species (143). The G4 structure in the promoter of c-*MYC*, if stabilized, can suppress the expression of this gene (144). On the other hand, the presence of a G4 structure in the promoter of the *OCT*4 gene increases the expression levels of this gene (145), rendering the understanding of the link between G4 structures and the expression level of the genes more complex than initially thought.

In immunology, the role of G4 structures is also very volatile. On the one hand, G4 structures allow pathogens to avoid immune detection and surveillance by facilitating DNA recombination at some specific loci such as in the *Pilin* locus and thus promoting antigenic variation (Av) (146–148). On the other hand, G4 structures binding by the LR1 protein, a DNA binding factor specific to B-cells, is a prerequisite for immunoglobin class switching (149).

Using a cell-free translation system, scientists theorized that G4 in the 5' UTR of RNA could modulate translation (150), but the mechanism by which the G4 regulates translation was not discovered in this study. One explanation of this phenomenon came from the study of the fragile X mental retardation protein (FMRP) that can also recognize a G4 structure in the 5' UTR of the SMNDC1's RNA affecting its translation by blocking the access of the translation machinery to the RNA (151).

1.1.2.3.1 Role of G4 structures in human diseases

Since we established that G4 structures could induce genomic instability and directed mutation and since these structures are present at key loci in the genome like the telomeres, origin of replication and key genes, it is only logical that these structures will be found to be implicated in human disease. In fact, many human diseases are a result of mutations in specialized G4 helicases. The WRN syndrome (152,153) described by Dr. Otto in 1904 is caused by a mutation of the WRN helicase (154), this helicase, among other roles, is specialized in G4 unwinding. This mutation leads to genomic instability since the telomeres of these patients are dysfunctional (155). Another disease that is a direct result of a mutation in helicases is the Xeroderma pigmentosum (XP) (156), this protein is also specialized in unwinding G4 structures among other roles.. XP presents the same telomeric genome instability as the WRN syndrome (157).

Cancer is a disease that also harbors a lot of genomic instability (158,159). Some of this instability can be explained by the errors in replication occurring at G4 structures while replicating the leading strand of the DNA (160). Many publications showed that the secondary structure of the DNA, especially the presence of G4 structures, is the leading reason behind the genome instability in cancer (161,162). In fact, this genome instability can be rescued by expressing the helicases

that unwind G4 structures in such cancer (163) since the helicases allow DNA replication through these structures (164).

G4 structures also affect neurons, especially at the *C9ORF72* loci linked to frontotemporal dementia (FTD) and ALS (165). The *C9ORF72* loci harbor a GGGGCC that can form a G4 structure with neuroprotective effects (166–168). In fact, the loss of TDP-43 in neurons, a G4 structure binding protein (169), can cause neuronal loss (170,171).

1.1.2.3.2 Using G4 as a therapeutic tool

Given the implication of G4 structures in many human diseases and ever since the discovery of molecules that can interact with these structures and stabilize them (142,172), many have searched the possibility of targeting these structures as a therapeutic means. This approach is prevalent in cancer treatment as a means to create a synthetic lethality in highly dividing cells by stabilizing the G4 structures (173,174), increasing the genomic instability in these cells. This approach is even more powerful when the cancer is lacking a G4 helicase, like the case of ATRX deficient glioblastoma (175). G4 stabilization in cancer can also be used as a complementary therapy along with traditional ones since it was shown that stabilizing these structures can increase the sensitivity of glioblastoma and lung cancer cells to radiotherapy (176,177). We should also remember that G4 structures are present and control the expression of many tumorigenic genes like *C-MYC*, *C-KIT*, *CD133* (178), and many others, this is another reason for the usage of G4 ligands to be added in cancer therapy (179). For example, CX-5461 G4 ligand was successfully used in cancer with *BRCA1* or *BRCA2* deficiency (180,181).

1.2 Chromatin compaction and polycomb complexes

Now that we understand the basic structure of the DNA, it is important to know some basic numbers. According to the latest genome compilation (GRCh38) we have 3,101,788,170 bp (182) in each cell spaced by around 3.4 Å, making the full length of an uncoiled DNA fiber in a cell around 1m [bp X vertical raise per bp = $(3 \times 10^{9})(0.34m \times 10^{-9}) = 1m$]. This 1m of DNA is packed in a nucleus of 6μm (6×10^{-6}m) of diameter on average (183). A simple calculation reveals that the circumference of the nucleus is around 19 μm which means that the DNA fiber can circle the nucleus 53 000 times. For this fiber to fit in the nucleus, it should be very condensed. In fact, in 1880, Walther Flemming first described very condensed parts of the nucleus that absorb dyes, thus baptizing it chromatin from the Greek word “chroma” or color (184). In 1951, Stedman described this dense structure as a mix of DNA and histones (185). Histones are “peptone like components of the cell nucleus” that can bind tightly to the DNA described by Kossel in 1884 (186). The histones act as scaffolds to contain these long DNA fibers in a specific architecture, compacting it to fit the nucleus. What is known as core histones, H2A, H2B, H3, and H4 form an octamer, two octamers are needed to form the nucleosome core particle, wrapping 147 base pairs of DNA around them (187). The nucleosome represents a higher organization of the DNA called “beads on a string”, or the 11 nm fiber, by Ada and Don Olins in 1974 (188), this conformation represents the primary condensation of the DNA. Linker Histones, H1 and H5, are present between these nucleosomes to compact them even more in what is called a 30 nm chromatin fiber (189–191). The 30 nm fiber represents the secondary condensation of the DNA fiber, and the tertiary condensation of the DNA fiber is the compaction of multiple DNA polymers via the interaction in a complex architecture (192).

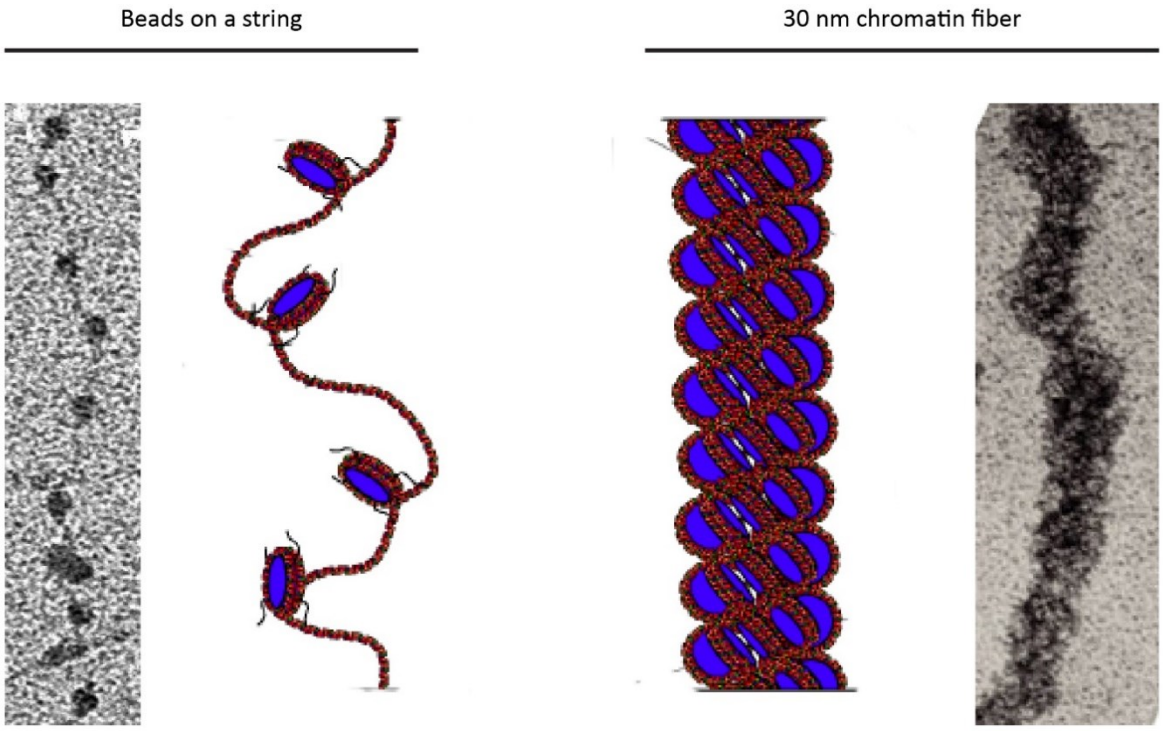


Figure 4. Structure of coiled chromatin (modified from (193–195))

DNA fibers are negatively charged due to the phosphate-sugar backbone (196), so from an electrostatic point of view, the DNA fiber should repulse itself making it unwind. In order to coil this negatively charged fiber, energy in the form of positive charges is required. In the chromatin, this energy is rendered by the histones since their tails are positively charged (197–199). This positively charged tail helps neutralize the DNA/DNA repulsion and helps to bend the DNA fiber in a compacted architecture (200).

1.2.1 Chromatin state

Emil Heitz, in 1928 reexamined the chromatin and noticed differences in the dye dependent on the cell cycle. This work was the first classification of the chromatin: heterochromatin or the chromatin that remains visible throughout the cell cycle, from the Greek heteros: other and chroma: color. The euchromatin is regions of the chromatin that loses the dye after mitosis, from the Greek eu: true and chroma: color (201). In 1966, Brown also divided the heterochromatin into two groups: constitutive heterochromatin and facultative heterochromatin (202), this distinction was at the time mainly based of the chromatization of the X chromosome (203,204). Since as

mentioned before, the positively charged tails of the histone are responsible for the compaction of the DNA, it is only logical that modifications to these tails are accountable for the degree of that compaction. In fact, the tails of the histones are particularly rich in residues that can be modified post-translationally. The most studied modifications are listed below:

Modification	Residue affected
Methylation (205)	Lysine, Arginine
Acetylation (205)	Lysine
Ubiquitination (206)	Lysine
Phosphorylation (207,208)	Threonine, Serine
ADP-ribosylation (209)	Glutamate
SUMOylation (210)	lysine

Table 2 . Post-translational modifications of the histone

These modifications are dynamic and reversible with enzymes that can deposit or remove them named “writers and erasers” respectively. A writer that can acetylate a residue is called acetyltransferase, the eraser of this modification is called deacetylase. A writer that can methylate a residue is called methyltransferase, and its eraser is labeled demethylase. And finally, a writer that can phosphorylate a residue is a kinase. In figures 5 and 6, reproduced courtesy of Cell Signaling Technology, Inc, we can see the vast numbers of these enzymes along with their target residues.

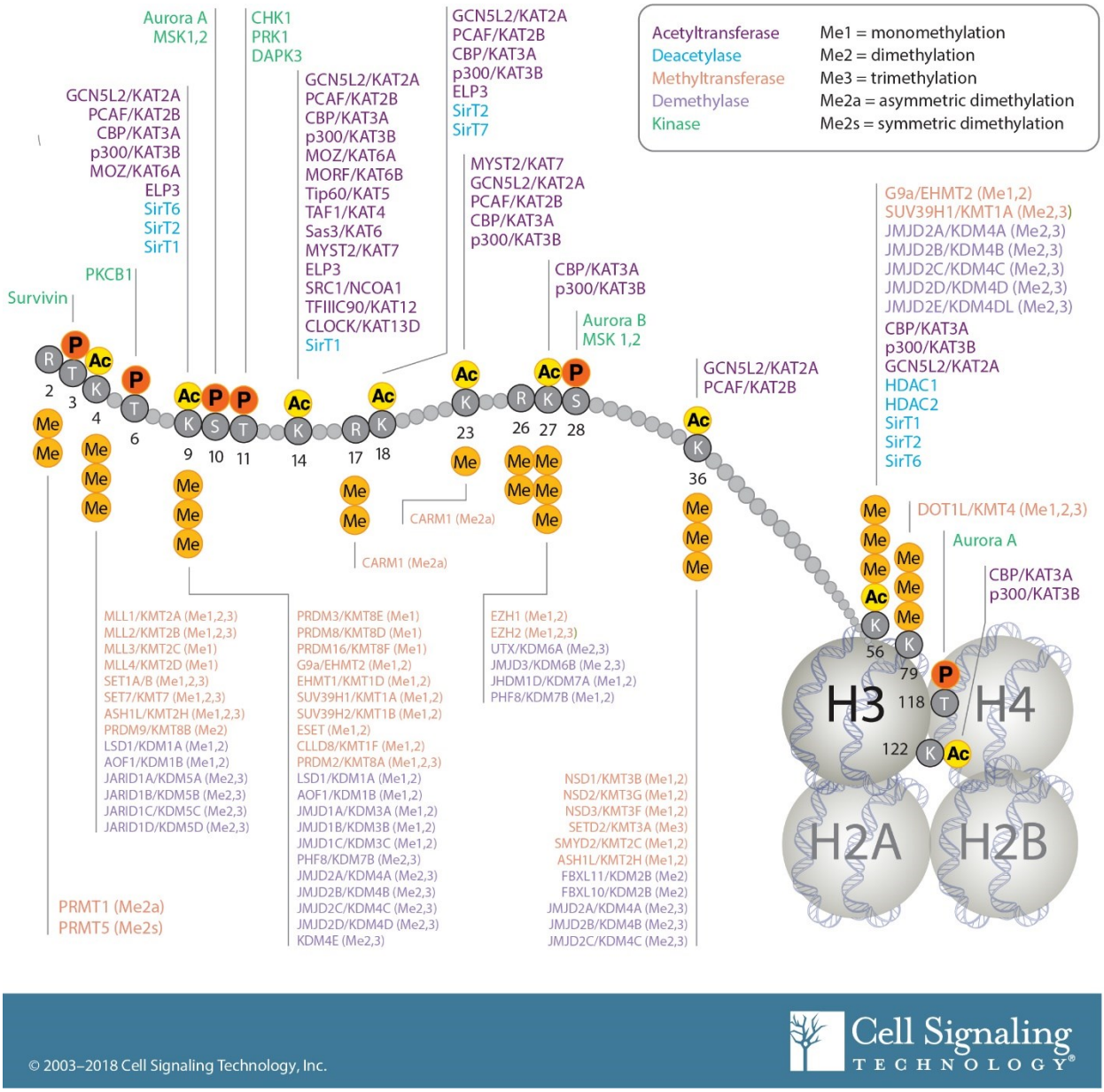


Figure 5. Writers and erasers of the histone's modifications of H3

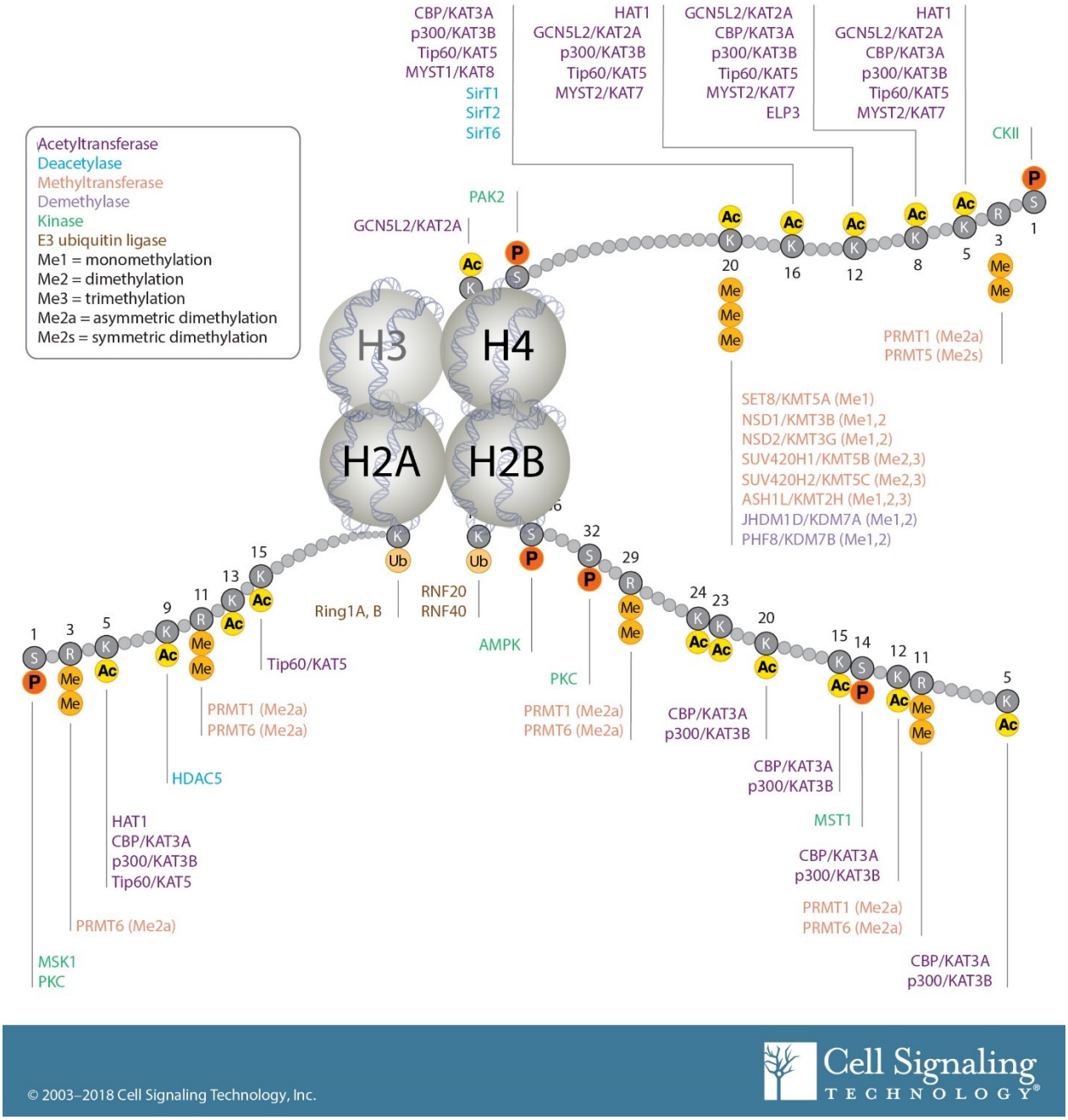


Figure 6. Writers and erasers of the histone's modifications of H2A, H2B, and H4

Some of these modifications change the electrostatic properties of the histone tails modifying the compaction of the chromatin, and thus the state of the chromatin, switching for example from closed facultative chromatin to loosely open euchromatin. All these modifications constitute what Waddington termed "the epigenotype" (211) in 1942. Waddington defined this domain as "the branch of biology which studies the causal interactions between genes and their products

which bring the phenotype into being." (211) Alternatively, the link that bridges between the genotype and the phenotype. This definition was changed over time, and with the advent of new technologies, nowadays epigenetics is defined as "the study of changes in gene function that are mitotically and/or meiotically heritable and that do not entail a change in the DNA sequence." (212). In simpler terms, epigenetics is the study of the effect of histone modifications on gene expression. Epigenetics also include DNA methylation, chromatin remodeling, histone variants (213) and non coding RNA (214). DNA methylation is the process by which a methyl group is added to the cytosine base (215,216), adenine base can also be methylated but this methylation is less studied (217). DNA methylation is primary performed by two DNA methyltransferase proteins called DNMT1 and DNMT3a/b (218,219). DNMT1 preferentially recognizes the hemimethylated DNA In mammalian cells (220). Hemimethylated DNA is formed during replication when the two strands of DNA are separated and copied, the nascent strand, or the newly formed, is unmethylated rendering the newly assembled two strands molecule hemimethylated. DNMT1 is then essential to copy the DNA methylation pattern from one strand to the other, that is why it is called a maintenance methyltransferase (221–223). On another hand, DNMT3a/b methylate during development previously unmethylated DNA, this is called "de novo" methylation (224–227). When DNA methylation occurs at regions rich in C and G bases called CpG islands present in the promoter of genes (228), it hinders the transcription of the related genes (229). In 1964, Allfrey was the first to describe such a correlation between histone modification and gene expression. In his article, Allfrey explains how histone acetylation allows the translation via RNA-polymerase, and he suggests that this acetylation may be a switch regulating gene expression (205). Since the time of Allfrey, and with the advent of ChIP technologies, we now know more and more about this epigenetic code and how each modification affects transcription. In table 3, we will summarize the effect of the most studied modifications in epigenetics.

	H3 K4	H3 K9	H3 K14	H3 K18	H3 K23	H3 K27	H3 K36	H3 K56	H3 K79	H4 K5	H4 K8	H4 K12	H4 K16	H4 K20	H2A K5	H2B K5	H2B K20
acetylation	(230,231)	(230)	(230)	(230)	(230)	(230,232)	(230,233)	(234,235)		(236,237)	(238)	(239,240)	(241)		(230)	(230)	(230)
mono-methylation	(230,242)	(230)		(230)	(230)	(230)			(230,243)			(244)		(230)		(230)	
di-methylation	(230,245)	(230)				(230)			(230,243)					(230)			
tri-methylation	(230,246)	(230)	(247)			(230)	(230)		(230,243)					(230)		(248)	

activation Activation/inhibition inhibition

Table 3 . List of the most studied histone modifications and their effect on gene expression

In the following section, we will describe these dynamic states of the chromatin, explaining the differences and the roles of each one.

1.2.1.1 Euchromatin

The euchromatin contains the actively transcribed genes of the cell. In fact, not more than 5% of our genome is actively transcribed in any given human cell (249). To be transcribed, the chromatin should be decompacted into the 11 nm fiber conformation in order to be accessible by the transcription machinery (250).

1.2.1.2 Facultative heterochromatin

The facultative heterochromatin is a designation for genomic regions containing genes that can adopt closed or open chromatin conformation, making them silent or transcribed, respectively. The choice of open or closed chromatin is dictated by the cell lineage and its maturation stage (251,252). In simpler terms, facultative heterochromatin contains, for example, a gene that should be expressed in early development but should be repressed in mature cells, or vice versa (253,254). To have this flexibility in gene expression, affected genes are decorated with bivalent

heterochromatin marks. H3K4me3 and H3K27me3, which are associated with gene activation and repression respectively, are bivalent marks that render the gene in a “poised” state that can be switched on and off dependent on the context (255).

1.2.1.3 Constitutive heterochromatin

In the literature, constitutive heterochromatin represents the genomic regions that are deserted from coding genes. These regions should always be silenced; that is why, in 1972, Ohno termed these regions as “Junk DNA” (256), it is evaluated that these sequences represent two-third of the human genome (257). Constitutive heterochromatin is usually in the 30 nm fiber conformation or higher conformation (258). The constitutive heterochromatin is enriched with transposable elements covering around 45% of the human genome (259).Furthermore, the heterochromatin contains vestige retroviral sequences integrated during evolution and maintained through it (260,261). These retroviral sequences represent, in a very conservative analysis, 8% of the human genome (262). This “junk DNA” if transcribed has the potential of reintegration in the genome, that is why they are also called “transposon elements” or “jumping DNA” (263). This reintegration can be very deleterious because it can target, for example, oncogenic genes inducing mutagenesis or activation, like the case of c-*MYC* (264,265) or *BRCA1/BRCA2* (266,267). The activation of these genes can be, in part, explained by the fact that these transposon elements harbor many transcription factor binding sites that can recruit the transcription machinery and induce the expression of these said genes (268).

Another role for heterochromatin is maintaining chromosomal integrity. This is because telomeres and centromeres are protected by this constitutive heterochromatin and represent a significant fraction of it. Telomeres are the extremities of the chromosomes, acting as a cap in order to protect the erosion of the chromosomes due to replication (269–273). On another hand, centromeres are regions in the core of the chromosome that act as anchors for sister chromatids pairing and control their segregation during mitosis (274,275). In fact, heterochromatin by the action of heterochromatin protein 1 (HP1) prevents telomere fusion. Overexpression of HP1 in human cells reduces the amount of telomere fusion along with the 3’ overhang, and increases the radioresistance of cells (276,277). It also maintains chromosomal integrity by reducing chromosome breakage and rearrangement (278,279), this is done by the role of heterochromatin

in the proper recruitment of the protein cohesion to the centromere safeguarding adequate chromosome segregation (280,281).

Furthermore, and like the mechanism described in the A-DNA, the compaction of the chromatin confers the genome refractory protection against DNA damage. In fact, the compaction of chromatin and the proteins involved in this compaction acts as a physical barrier to DNA damage reducing the damage dealt from radiation (282,283) as well as any damage from reactive oxidative species (ROS) (284–286).

While the primary role of the heterochromatin is believed to be the protection of the genome (287–295), and while it is deserted from genes, heterochromatin plays a significant role in gene expression. This regulation of gene expression is made possible by the reorganization of nuclear domains, where genomic regions are packed in separate geographic regions of the nuclei called topologically associated domains (TAD) (296–298). In simpler terms, the 3D organization of the nucleus can influence gene expression by trans or cis regulation.

1.2.1.3.1 Silencing in trans

The 3D organization of the genome inside the nucleus is proving to be very important. In fact, Nuclear lamina is made of intermediate fibers called lamin that anchor the chromatin in the nucleus (299) with the heterochromatin at the periphery (201) and the euchromatin in another compartment of the nucleus (300). These compartments are called TAD, the compartment that harbors the anchoring of the heterochromatin to the lamina is called Lamina associated domain (LAD). The LADs are silent heterochromatin, while the other TADs usually contain active genes. Furthermore, highly expressed genes are organized in defined regions called regions of increased gene expression RIDGEs (301). During differentiation, genes can be shifted from active TADs to silenced LADs and, in effect silencing these genes (302). A visual explanation is presented below (Fig. 7), showing a red region being silenced in trans when it was reorganized and shifted toward the LAD.

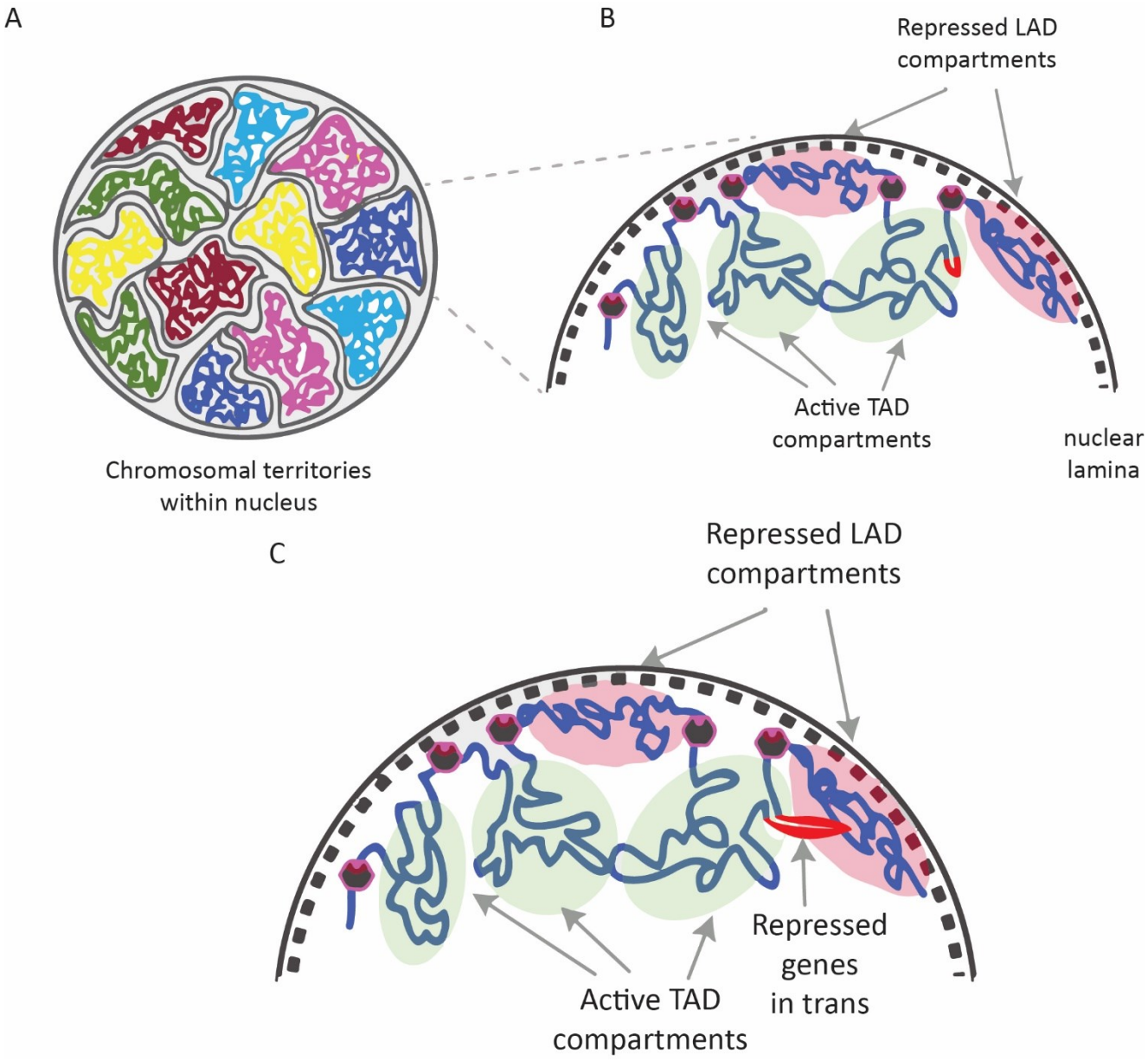


Figure 7. Gene silencing in trans by the heterochromatin (based on (303,304))

A. 3D representation of the nucleus with the delimitation of the chromosomal territories.
B. A magnification of the nuclear membrane showing the anchoring of the LADs along with the target region (in red) that is active
C. The target region (in red) was moved during differentiation putting it in the LAD compartment and thus effectively silencing it.

1.2.1.3.2 Silencing in cis

This type of silencing was first described in 1930 by Muller in the *Drosophila*. Muller named this type of silencing position-effect variegation (PEV) (305,306). PEV occurs when a part of the heterochromatin is translocated near a gene. This translocation does not induce mutation in the gene or its promoter, but it spread the heterochromatin domain beyond its normal reach, heterochromatizing the gene in question (307,308). While PEV was first described in *Drosophila*, this regulation also occurs in mice, inactivating some transgene (309,310), and some pieces of evidence are emerging that it also happens in human cells via the Human Silencing Hub (HUSH) complex (311).

It is essential to mention that while the definition of chromatin was based on the retention of dye at first, nowadays, we have a molecular signature that can distinguish between the three chromatin domains (see table 4).

	euchromatin	facultative heterochromatin	constitutive heterochromatin
DNA methylation	Hypomethylated		hypermethylated
associated proteins	H3.3	HP1γ	HP1α/β
histone marks	Hyperacetylation H3K9ac H3K4me2 H3K4me3 H3K36me3	Hypoacetylation H3K9me2 H4K20me1 H2AK119ub1 H3K27me3	Hypoacetylation H3K9me3 H4K20me3

Table 4 . Commonly used markers to distinguish between the various states of the chromatin (312)

1.2.2 Polycomb repressive complexes

Many proteins can write and erase epigenetic marks, as we saw in (Figs. 5 and 6). In this section, we will focus on one of the most influential writers' complexes, and that is the Polycomb-group proteins (PcG). This complex was discovered in 1978 by Lewis as a master switch controlling the development and the specification of cellular identity in *Drosophila* (313). In this study, Lewis described how the PcG is responsible for maintaining the repressive state of the *bithorax* gene complex during development. The *bithorax* gene complex is a part of the homeotic gene complex

of the *Drosophila* (314,315) regulating the development of anatomical structures (316). This complex was then proved to be conserved in mice (317,318), not just by protein homology but even with conserved functions (319). This complex was also evolutionary conserved in humans (320,321), proving the importance of such a complex. In mammals, PcG was not only linked to *HOX* gene silencing, but it was also linked to the proper differentiation of cells into their intended mature state controlling their fate from the embryonic stage till the adult one (322–324) by controlling the developmental regulators of stem cells (325,326). PcG was also involved in managing the senescence of cells (327–330), and chromosome X inactivation (331–333). All these discoveries point to the role of PcG in the formation and maintenance of facultative heterochromatin silencing genes dependent on the cell environment. Indeed PcG was found to be enriched at said facultative heterochromatin (334). Considering the importance of PcG, it was extensively studied and linked to many human diseases like cancer (335,336), and other development diseases (337). In humans, this complex is divided into two distinct complexes: the Polycomb repressive complex 2 (PRC2), and the Polycomb repressive complex 1 (PRC1).

1.2.2.1 PRC2

The primary role of the PRC2 complex is to initiate chromatin silencing by the trimethylation of H3 on its lysine in the 27^{th} position (H3K27me3). The repressive function of this complex is carried out by the EZH2 subunit, the ortholog of the *Drosophila* E(z) protein (338). EZH2 trimethylates H3K27 via its SET domain (339,340). SUZ12 protein, part of the same complex, is essential for this activity and the stability of the complex since it's knockout decreases the level of EZH2 (341). Moreover, EZH2 on itself cannot function as a methyltransferase, it should be associated with SUZ12 and EED in order to tri- and di- methylate H3K27 (342–345), however, only EED is required for the monomethylation of H3K27 (346). In figure 8, we can see the various core components of the mammalian PRC2 complex and their interactions with the chromatin. The two components: RbAp46 and RbAp48, also called P55, act as a chaperone that binds the histone and anchors the complex to its place (347). JARID is an essential part of the PRC2 complex, especially during the embryonic stage, regulating the differentiation of stem cells into mature cells (348,349). It is crucial to mention that PRC2 methylates itself, specifically JARID, to gain the function of methylating the H3K27 (350).

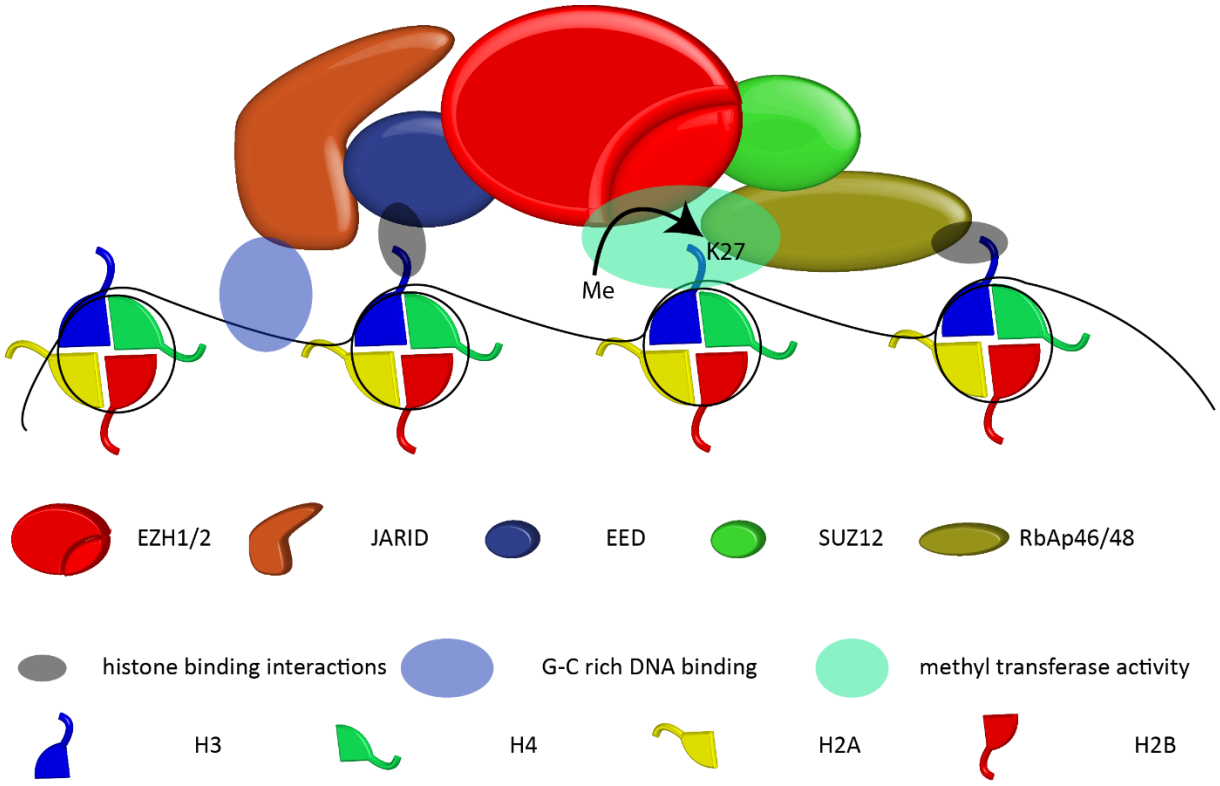


Figure 8. PRC2 complex and it's interaction with the chromatin (351)

In summary, the PCR2 complex is evolutionarily conserved from its ortholog in *Drosophila*, maintaining all the domains essential for its role, yet expending the number of proteins with homologous counterparts (352). In table 5, we can see a comparison of these two complexes and the proteins involved in their functions.

PRC2

domain	*Drosophila proteins*	mammals' proteins	function
SET domain	E(Z)	EZH1	H3K27 methyltransferase Histone binding
SANT domain		EZH2	
Zinc-finger domain	SU(Z)12	SUZ12	Possible DNA/RNA binding
	Jing	AEBP2	H2AK119Ub1 binding
WD40-repeat domain	ESC	EED	H3K27me1/2/3 binding
WD40 domain	NURF55	RbAp46	H3K36me3 binding
		RbAp48	
PHD-finger domain Tudor domain	Pcl	PCL1	H3K36me3 binding
		PCL2	
		PCL3	
ARID domain		JARID2	RNA binding

Table 5 . PRC2 comparison between the *Drosophila* and the mammalian complex (337)

1.2.2.2 PRC1

In the most referenced pathway, PRC1 is recruited to the chromatin by the PRC2. PRC1 contains a chromobox protein (CBX) that can read the H3K27me3 mark deposited by the PRC2 and recruit the PRC1 complex to this mark (353,354). Once recruited, PRC1 by the action of RING1B and BMI1 ubiquitinates H2A on its lysine in the 119th position (331,355) see figure 9 for model representation. Furthermore, PRC1 was shown to interact with another histone modifier: the SUV39H1 (356). SUV39H1 contains two structural domains: a chromodomain and a SET domain (357). Via these two domains SUV39H1 plays a central role in heterochromatin spreading (358): SUV39H1 is recruited via the chromodomain that recognizes the methylation of the H3K9 (359), then the SET domain is responsible for the histone methyl-transferase activity, methylating surrounding histone H3 at their lysine 9 (360). Since the chromodomain of SUV39H1 can recognize the product of this same protein this reaction is repeated spreading the heterochromatin around the initial site establishing a positive feedback loop (361,362). The H3K9me3 mark in then recognized by the chromodomain of HP1 (363,364). HP1 recruitment to the heterochromatin induces its conformational change allowing the spreading of the heterochromatin (365) in a bidirectional way (366).

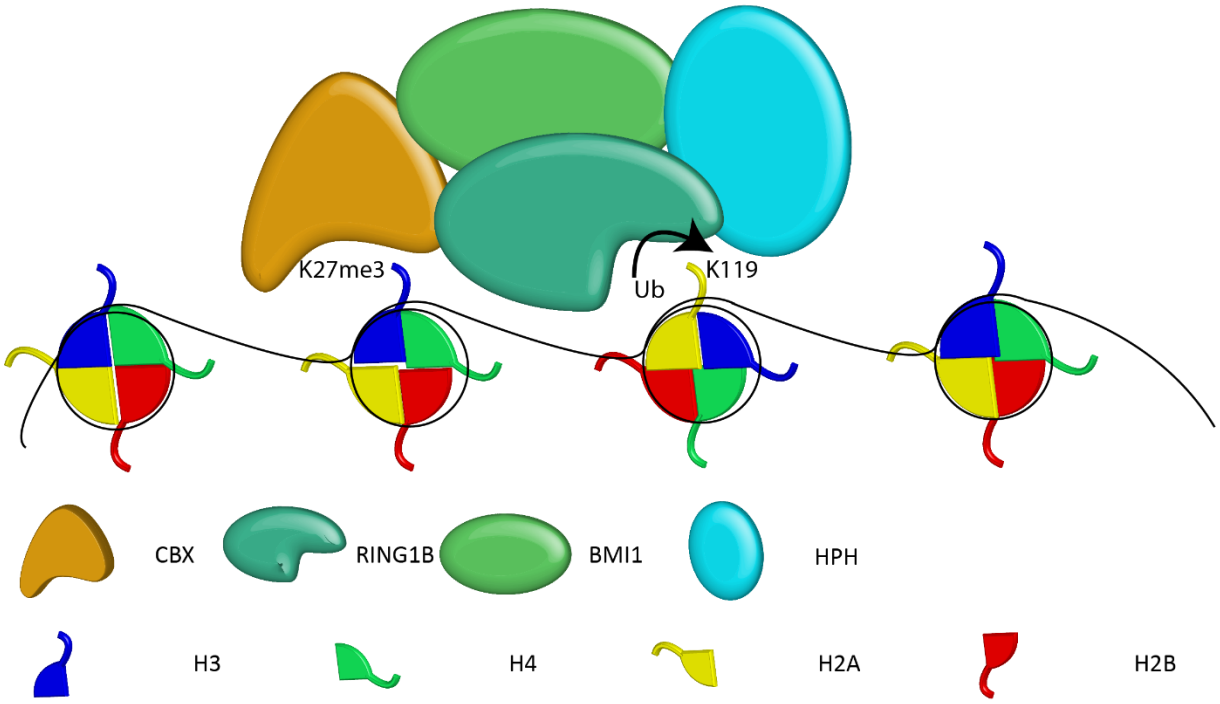


Figure 9. Recruitment and function of PRC1 in mammalian cells (367)

As any canonical pathway, the hierarchical recruitment of PRC2 then PRC1, the more we study it, the more we can discover that it is more diversified and more complex than initially thought (368). In fact, in embryonic stem cells lacking EED, a vital component of the PRC2 complex, H3K27me is absent, but PRC1 can still ubiquitinate its target genes showing that PRC1 can act independently of PRC2 and be recruited by other mechanisms than the chromodomain of CBX (369). This activity is mediated by a variant of PRC1 that has its CBX replaced by the RING1 and YY1-binding protein (RYBP) (370,371). With some evidence showing direct recruitment of PRC1 to chromatin by transcription factors like the Runx1/CBFβ (372). Furthermore, new studies revealed that the recruitment of PRC2 could be dependent on the H2A ubiquitylation deposited by a variant of PRC1 inversing the canonical recruitment model that usually state PRC1 is recruited to PRC2 loci (373,374). More specifically, a variant of PRC1 mono-ubiquitylate H2Aub at lysine 119, this leads to the recruitment of PRC2 by the subunit JARID2 (373,375,376), leading to the silencing of PcG target genes (377). Other studies revealed that H2A ubiquitination by PRC1 is sufficient for the

gene repression of Polycomb target genes (378) and that without the involvement of PRC2. The story gets more complicated because some studies showed, that constitutive PRC1, and not variant PRC1, create a positive feedback loop on the PRC2 to reinforce and maintain gene repression. In fact, the loss of H2Aub mediated by the loss of PRC1 reduces the levels of H3K27me3, leading to higher gene expression in targeted genes (379,380).

Like the case of PRC2, the proteins of PRC1 were duplicated during evolution that is why in mammals we have *RING1A* and *RING1B* that are duplication of the same genes with a slight difference in roles (381) and we have *BMI1* and *Mel18* also a duplication of the same genes with slightly diverse functions. To summarize the diversification of this complex, please refer to table 6 for a comparison of this evolution.

PRC1

domain	*Drosophila* proteins	mammals' proteins	function
chromodomain	Pc	CBX2	H3K27me3 binding (all) H3K9me3 binding (CBX4) RNA binding (CBX4 and CBX7)
		CBX4	
		CBX6	
		CBX7	
		CBX8	
Sterile alpha motif (SPM domain)	Ph	PHC1	Possible RNA binding
		PHC2	
		PHC3	
	Scm	SCMH1	
		SCMH2	
RING-finger domain	dRING/Sce	RING1A	Histone Ubiquitination (H2AK119Ub1)
		RING 1B	
	Psc	NSPC (PCGF1)	
		MEL18 (PCGF2)	
		PCGF3	
		BMI1 (PCGF4)	
		PCGF5	
Zinc finger domain	dRYPB	RYPB	DNA binding
		YAF2	
		KDM2B	H3K36me3 demethylase unmethylated CpG island binding

Table 6 . Comparison table between the *Drosophila* PRC1 and its mammalian homologs (337)

1.2.2.2.1 BMI1

In this work, we are going to focus on *BMI1* or "B Lymphoma Mo-MLV insertion region 1 homolog", a protein part of the PRC1 complex. *BMI1* was first discovered in 1991 as an oncogene that cooperates with *Myc* in the lymphomagenesis of B lymphoid tumors in mice (382,383). This gene on Chromosome 10 and contains ten exons (Fig. 10-A). When translated, the final protein is 326 amino acids and weights 37 KDa. BMI1 is mainly localized in the nucleus of cells due to its nuclear localization signal domain (NLS) localized at between the 42 and the 45th position (Fig. 10-

B). Depending on the prediction model another NLS domain is also expected between the position 232 and 235 of BMI1 (384–386). Another essential feature of BMI1 is its RING finger domain, which binds the ubiquitination enzyme, in this case, RING1B, and it's substrate, in this case, H2A, and acts as a ligase (387,388). In fact, the formation of the complex BMI1/RING1 is essential for the ubiquitin ligase activity (389,390). This heterodimerization allows the autoubiquitination of RING1 (391). While this ubiquitination does not affect the ubiquitin ligase activity of the complex BMI1/RING1 in vitro, it does increase it in vivo, suggesting that this ubiquitination stabilizes the complex and reinforces its ubiquitin activity in vivo (392). In fact, the formation of the BMI1/RING1 complex reduces the affinity between RING1 and UbcH5c, the only ubiquitin ligase that can polyubiquitylate RING1 leading to its degradation, thus increasing the stability of the PRC1 complex (393). The C-terminal of BMI1 encapsulates a proline (P), glutamic acid (E), serine (S), and threonine (T) rich domain known as PEST domain (Fig. 10-B), that directs the protein degradation via the proteasome shortening the half-life of BMI1. The deletion of this domain extends the half-life of BMI1 and stabilizes it (386). Furthermore, BMI1 contains two domains surrounding a Helix turn Helix domain that can interact with PHC2 (394) and E4F1 (395), and a domain around the Ile212 that allows the homo-dimerization of BMI1 (394). All these interaction domains of BMI1 and its ability to self-dimerize makes BMI1 one of the central protein in the PRC1 complex that organizes its architecture by self-dimerizing and mediating its interaction with most of the other components like RING1B, PHC2, and E4F1 (394).

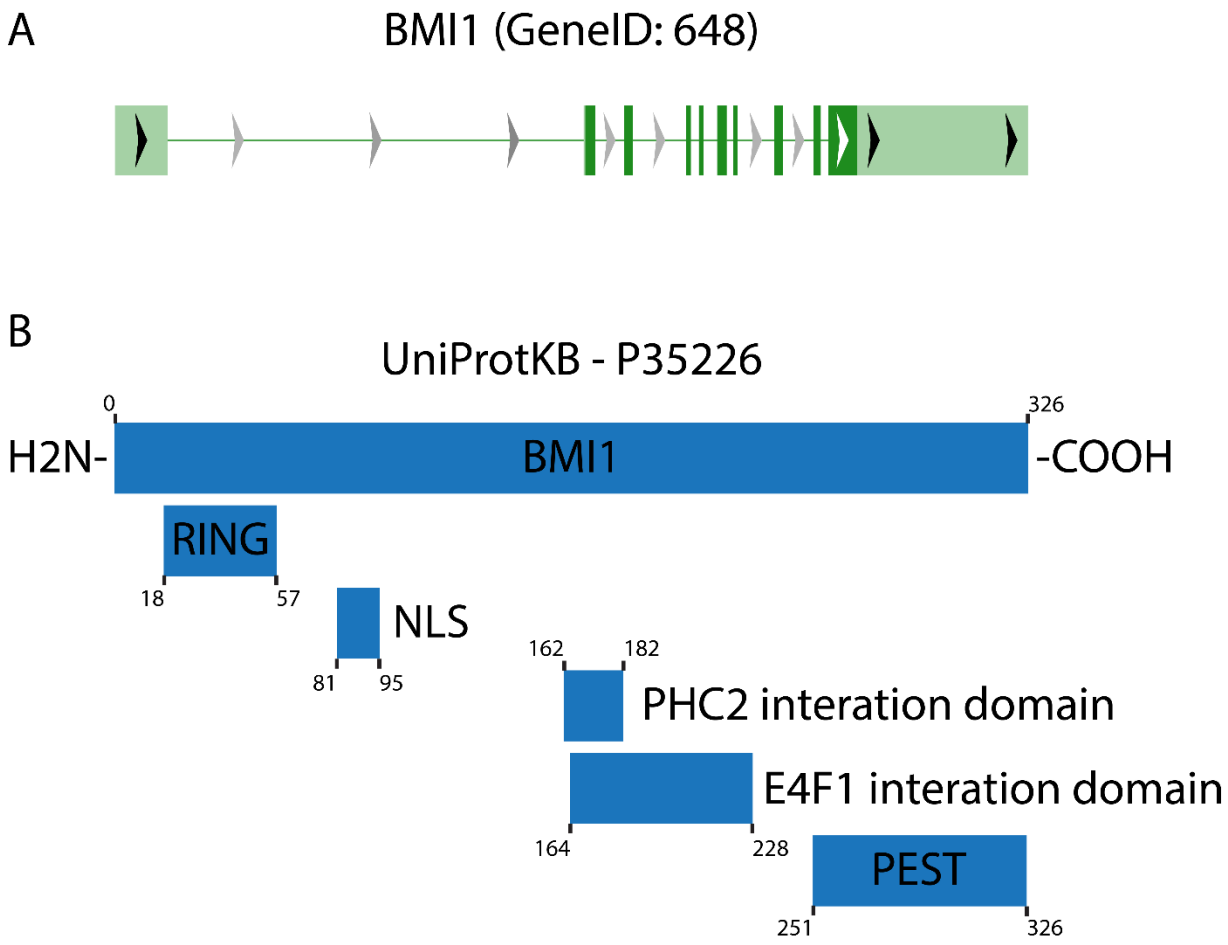


Figure 10. A visual representation of the *BMI1* gene and the BMI1 protein highlighting the structural domains.

A. A visual representation of the genomic location of *BMI1* (geneID: 648) showing the exons and the introns (396).
B. Visual representation of the BMI1 protein (P35266) as referenced in the Uniprot database. The first line shows the full-length protein, while the other lines show the structural domains of this protein along with their localization (397).

BMI1 is ubiquitously expressed and translated in all organs, as shown in figure 11 from the proteomic and genetic data collected by the Human Protein Atlas (398). However, BMI1 is overexpressed in many cancers, as shown by the Gepia database (399,400). In fact, BMI1 plays an essential role in the tumorigenesis of many cancers including prostate cancer (401–403),

pancreatic cancer (404), colorectal cancer (405–407), lung (408–410), nasopharyngeal carcinoma (411–413), breast cancer (414,415), along with many central nervous system (CNS) cancers like medulloblastoma (416–419), neuroblastoma (420–422), and glioblastoma (423–426).

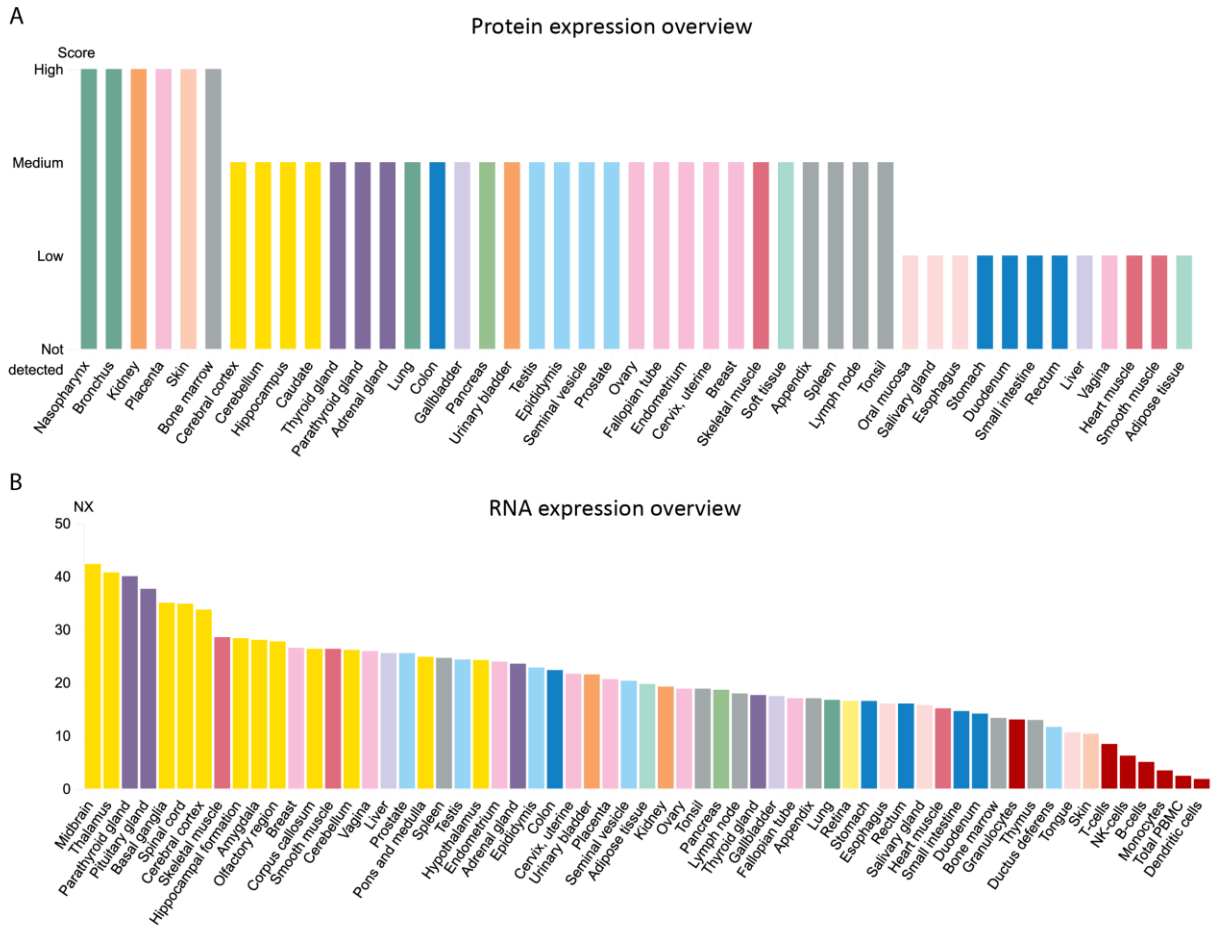


Figure 11. Protein and RNA expression of BMI1 across various human tissues according to the human protein atlas (398).

A. Protein expression data from the human protein atlas over 44 tissue is shown in a bar graph
B. RNA expression data build by the human protein atlas based on three databases: Human Protein Atlas (HPA) RNA-seq data, RNA-seq data from the Genotype-Tissue Expression (GTEx) project and CAGE data from FANTOM5 project.

BMI1 broad implication in cancer progression is not surprising since BMI1 overexpression is sufficient to immortalize cells by activating the human telomerase reverse transcriptase (hTERT) expression (327,427–429). Alternatively, BMI1 can prevent senescence and apoptosis by silencing its key target, *Ink4a/Arf* locus. This locus encodes for $P16^{Ink4A}$ and $P19^{Arf}$ (430). In normal conditions, $P16^{Ink4A}$ prevents the association between cyclin D and cyclin-dependent kinase 4 and 6 complexes (CDK4/6) (431). If these two proteins are not associated, they cannot phosphorylate the retinoblastoma protein (Rb) in its C terminal (432), thus allowing it to sequester the transcription factor E2F (433). In the case of BMI1 overexpression, P16 expression is silenced. Thus, cyclin D and Cdk4/6 can be associated, and hyper phosphorylate Rb. Hyperphosphorylated Rb cannot bind E2F, allowing it to initiate the transcription of genes essential for G1/S transition (Fig. 12). BMI1 also repress *P19*, in its turn P19 sequesters mouse double minute 2 (MDM2) protein, a ubiquitin ligase that targets P53 and induces its proteasomal degradation (434,435), so in BMI1 overexpression, P53 is ubiquitylated decreasing its half-life and its stability (436) (Fig. 12). Furthermore, BMI1 and RING1 can directly bind P53 and ubiquitinate it, leading to its degradation and the promotion of cancer cell proliferation (437), especially in cancers affecting the CNS like the neuroblastoma and medulloblastoma (436). Since these pathways are critical in tumorigenesis, they are often targeted in cancer therapies (438–440).

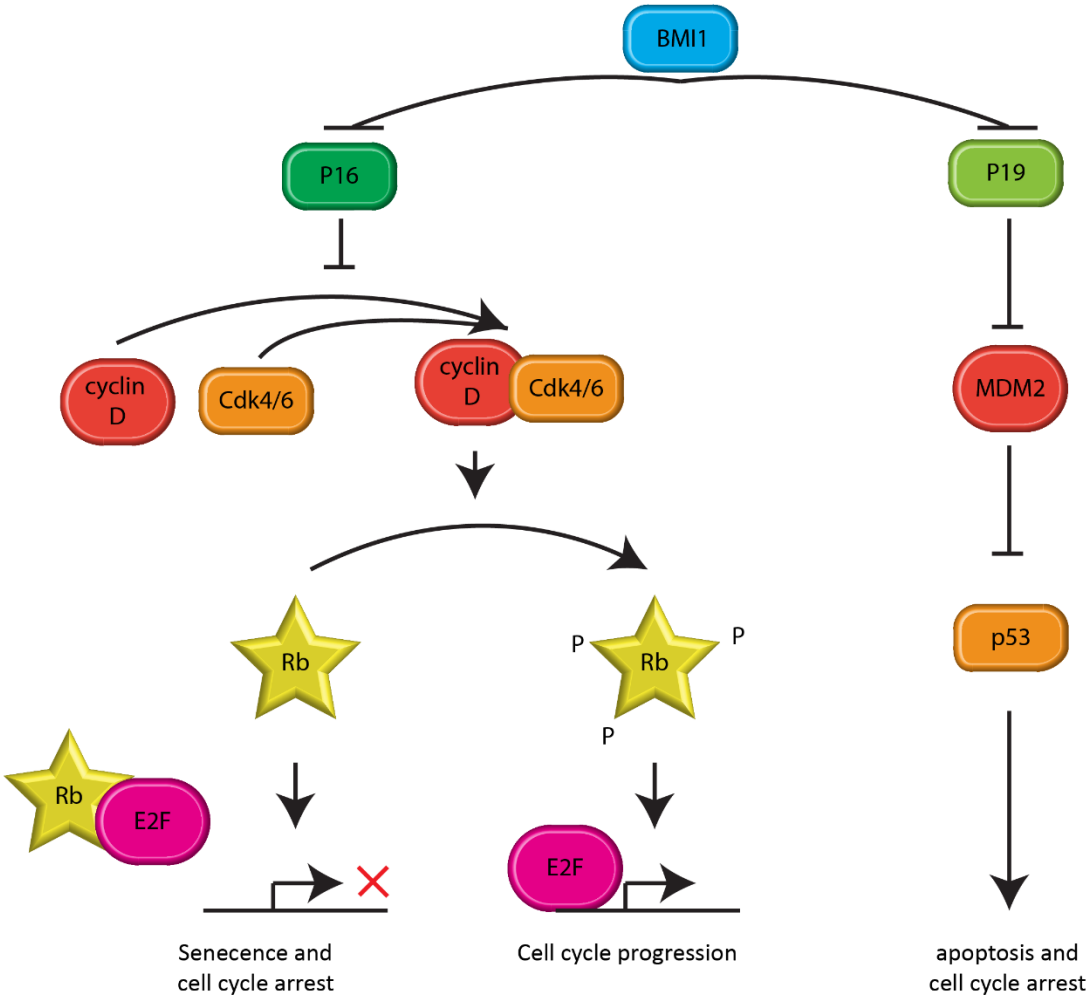


Figure 12. BMI1 regulates cell cycle via P16 P19 pathways

1.2.2.2.1.1 BMI1 in the CNS

While BMI1 is ubiquitously expressed in the organs, many studies have highlighted its role in the CNS. BMI1 null mice have development problems inducing a defect in cerebellar growth (441). These mice exhibit neurological problems at birth, like ataxia, epilepsy, and generalized astrogliosis (442). The development problems of these mice are not restricted to the cortical neurons; in fact, these mice exhibit a retinal development problem as well, and the cone photoreceptors start degeneration soon after birth (443).

BMI1 was shown to be reduced with aging in humans and mice (285,444,445). This reduction induces the aging phenotype (446) and can favor the onset of age-related disease. Recently the

reduction of Bmi1 in mice was associated with neurodegeneration and Alzheimer like phenotypes (447). In fact, in sporadic Alzheimer disease patients (sAD), BMI1 was reduced, and the reduction of BMI1 in healthy cortical neurons was sufficient to induce AD phenotype in culture (448). The role of BMI1 as a neuroprotector is often associated with its role in the activation of antioxidant response genes (AOR). Under oxidative stress NRF2 induces the transcription of these AOR, and P53 acts as an inhibitor of NRF2 reducing the transcription of AOR. BMI1 by degrading P53 can reestablish the expression of these genes, thus reducing the harmful reactive oxygen species (ROS) in the neurons (285,286). Still, recent studies have also demonstrated that BMI1 plays a crucial role in repressing *MAPT* gene or microtubule-associated protein tau (this gene codes for the Tau protein). Hence, the loss of BMI1 in neurons leads to the expression of *MAPT*, therefore an increase in Tau proteins (448). Furthermore, loss of BMI1 stabilizes GSK3b and P53, increasing the accumulation of Tau proteins and decreasing the proteasomal degradation of these proteins (448).

1.2.3 Repetitive DNA

Previously we mentioned that one of the roles of constitutive heterochromatin was to maintain genome stability by repressing repetitive sequences of the genome that are overrepresented in it (449). In fact, while these repeats represent 45% of the human genome (259), the underlying genomic sequence of these repeats is less important for genome stability than the chromatin conformation surrounding them and the maintenance of that chromatin (450). DNA repeats are considered in fact “Fragile sites” (451) prone to mutation, number variation, chromosomal breaks, and replication stalling or arrest (452–454). On the other hand, if these sequences are well maintained and structured by the chromatin, they can be a potent tool for the cell to ensure its chromosomal stability (455), especially at telomeres (456). These structures can also help the cell control its gene expression (457) and even play a significant role in evolution by favoring gene duplication and mutation (458–460). These elements can be categorized in many ways; in the next paragraphs, we will divide these elements into two major groups: the tandem repeats and the interspersed repeats explaining the differences. In figure 13, we can see a summarized graph of the repeats type in the human genome with their corresponding families. The data was compiled from various publications along with the RepeatMasker database (461).

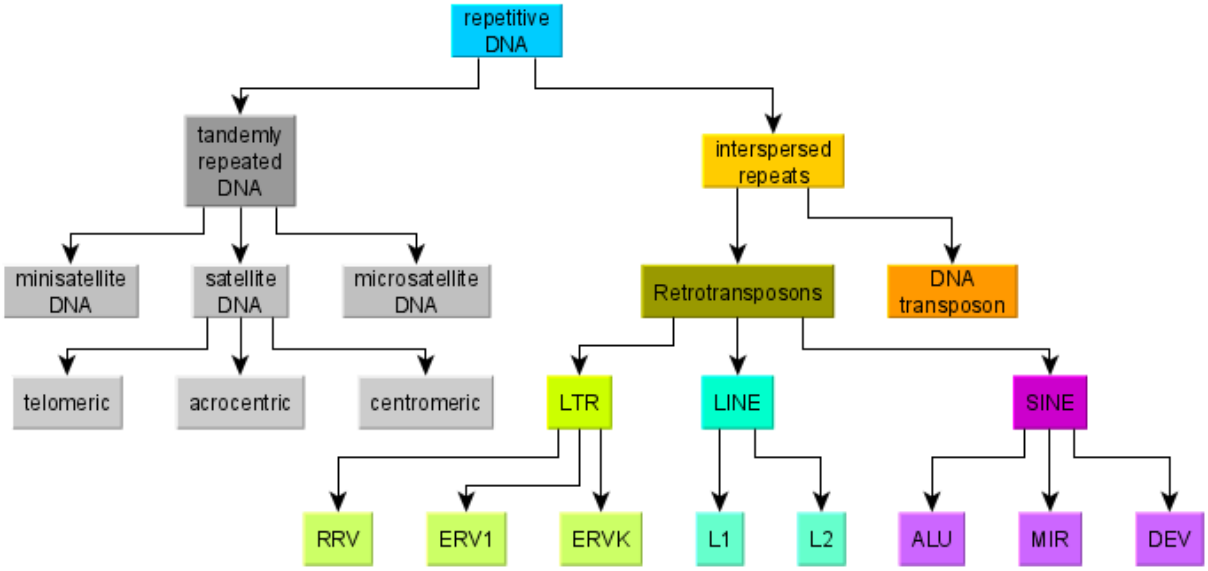


Figure 13. Hierarchical classification of DNA repeats

1.2.3.1 Tandem repeats

Tandem repeats are highly repetitive sequences, usually repeated in tandem without any non-repetitive sequence interspersed between them (462). The minisatellite DNA is formed of stretches of 9 to 100 bp repeated in clusters of 10 to 100 while the microsatellite also called short tandem repeats (STR) are stretched of 3 to 5 bp repeated in clusters of 10 to 100 (463). The primary family of this category is the satellite DNA that englobes the acrocentric, the centromeric, and the telomeric DNA according to the RepeatMasker database (Fig. 13) (461).

Tandem repeats are, by nature, very prone to what is called “replication slippage” or slipped-strand mispairing (464). We have discussed the concept of SSM and its implication on genome integrity previously. In the context of tandem repeats, an SSM event induces a copy number variation within the cluster, turning a proto-microsatellite, for example, into a microsatellite (465). Satellite DNA can also be beneficial, especially during mitosis, when satellite DNA has a unique role in the pairing and the segregation of sister chromatins (466–468). Satellite DNA also plays a crucial role in telomere capping (469).

1.2.3.2 Interspersed repeats

The building blocks of the interspersed repeats are the transposon, or the part of the genome that can jump, and migrate from one locus to the other duplicating itself. Transposons were discovered in 1940 by Barbara McClintock, a discovery that earned her a Nobel prize in 1983 (470). These transposons or mobile elements can be duplicated and spread across various loci on the genome creating the interspersed repeats.

The interspersed repeats can be divided into two categories depending on the mechanism underlying the transposition of these elements: some transposon jump from one locus to the other in the form of DNA using a protein called transposase to facilitate their integration in the target locus. These transposons are called DNA transposons (Fig. 13). In this case, the template sequence is cut from the original locus, or replicated then cut, and it travels to the target locus to be integrated (471). Other transposon elements require to be transcribed, thus passing by an RNA transitional phase before being incorporated back into the genome by a retrotranscriptase, these elements are called RNA transposons or retrotransposons (472).

From these two categories of interspersed repeats, only the retrotransposons are considered to be active and functional in the human genome nowadays (259). The retrotransposons can be further divided into LINE or SINE, depending on the length of the transposon element.

1.2.3.2.1 LINE

LINE elements or Long interspersed nuclear elements, as their name indicated, are long elements usually of 6 to 7 kb. Only the most conserved of the LINE elements still harbor all the elements and code for all the proteins needed for transposing efficiently (473). The proteins required for the retrotransposon machinery are under two open reading frames (ORF) in the LINE elements: *ORF1* code for the P40 coiled coil-mediated trimeric proteins that play as a chaperon for packaging the RNA (474–476), this P40 protein is similar to the transposase 22 protein (474). *ORF2* code for the ORF2p protein that has a reverse transcriptase activity and an endonuclease activity, it is this endonuclease activity, precisely it's capacity to cleave the DNA, that is responsible for the reintegration of the LINE element into the genome. (477–480). For example, the insertion of L1 elements near the *NF1* gene causes a disease called neurofibromatosis (481). LINE elements,

specifically the L1 family, are highly active during embryonic neuronal differentiation (482). The retrotransposition of these L1 elements in somatic cells may also be beneficial, creating diversification within the same organism, for example, between the different neuronal cells. This diversification in the genetic load of neurons was hypothesized to be important to increase the neuronal plasticity (483).

1.2.3.2.2 SINE

SINE elements or Short interspersed nuclear elements, as their name indicated, are short elements usually of around 0.3 kb. Unlike LINEs, SINEs are non-autonomous and required the transposon machinery of LINEs to jump loci, that is why they are considered parasitic elements (484,485). When these elements are inserted in or near genes they can cause diseases, in fact, their insertion was associated with cancer (267), hemophilia A and B (486–488), cystic fibrosis (489), and neurofibromatosis (481).

1.3 Ageing, Alzheimer disease and neurodegeneration

In this last section of the introduction, we will be talking about the concept of aging, first from a cellular point of view to the full organism's point of view, highlighting the neurodegeneration that accompanies the aging of humans. The concept of aging has shifted a lot in the history of science along with the concept of immortality, and nowadays we are experiencing another major shift in this domain with an increased focus on how to become “immortal” or more precisely how to extend the life span of humans (490). In order to fulfill this task, some scientists are promoting a change in lifestyle (491), while others are relying more on drugs (492) to curb the aging process and slow it down. One of the main targets of such drugs is senescence (493). In this section, we will try to understand the aging process better and apprehend its significance.

1.3.1 Hayflick limit

The dogma in the scientific community was that cells are immortal if cultured outside the organism and that only extracellular phenomena can affect the life span of cells. This dogma was mainly propagated by an experiment done by Dr. Alexis Carrel, who reported in 1912 the serial cultivation of cells derived from the embryonic chicken heart for more than 20 years (494). In 1961, while establishing a new cell line called WI-38 (495), Dr. Leonard Hayflick noticed that the cells were dying after 50 passages or what he called subcultivations (496). Dr. Hayflick called this phenomenon the “Phase III”, arguing that each cell had a limited internal number of divisions that it can accomplish. This new view of the cellular mortality introduced the concept that intracellular phenomena can also be responsible for cell senescence. It was Sir Macfarlane Burnett, a Nobel laureate, that coined the term “Hayflick limit” to this new way of thinking about cellular senescence in his book Intrinsic Mutagenesis: A Genetic Approach to Ageing (497). Cellular senescence, or more precisely in this context cellular replicative senescence, is an irreversible arrest of the cell cycle (498). Senescent cells are also distinguished by their secretomes that are high in growth factors, proteases, inflammatory cytokines, and immune modulators; all these secreted factors combined are called Senescence-associated Secretory Phenotype (499). In the next section, we will be introduced to some of the modulators of this Hayflick limit in cells and the molecular mechanism behind these modulators.

1.3.2 Aging of the cell

In order to better understand the aging of an organism, we ought to understand the aging process of its building blocks: the cells that form this organism, that is why we are going to explore some of the factors that influence the aging of these building block, the senescence of the cells, and its effect on the organism as a whole (500).

1.3.2.1 Telomere shortening

Dr. Hayflick continued his work on cellular senescence and established that there is an internal counter on the number of divisions that a cell can undergo; he called this the replicometer. Dr. Hayflick and Dr. Wright also proved that this replicometer was inside the nucleus (501,502). At the time, the molecular mechanism behind this replicometer was unknown. The more we understood the molecular aspect of biology, the more telomere became interesting candidates of this replicometer since with each round of cell division, these telomeres get shorter (503). It was proven that when these telomeres reach a certain critical length, the cell undergoes senescence (504–506). The experiment that consolidated the idea that telomere shortening is the replicometer is the introduction of telomerase in cells that lacked this protein. Telomerase protein helps to replicate the telomeres allowing them to regain length (507). The overexpression of this protein extended the life span of these cells showing the importance of telomere length on the senescence of the cells (508,509).

In order to better understand the importance of telomeres, we need to understand their functions. Telomeres are tandem repeats at the extremities of our chromosomes that span 5kb to 15 kb of the human chromosomes (510). This stretch of DNA ends with a G-rich leading strand that is in the single strand form called G-overhang. This G-overhang invades the C-rich double strand region that is upstream of it, forcing the end of the telomere to create a loop, this loop is called a T-loop (511,512). This structure is maintained and controlled by a complex called the shelterin complex (277,513). These loops, along with the shelterin complex, are essential to preserving the integrity of the chromosome by preventing the fusion of the telomeres at their respective ends (514). They also play a role in circumventing the DDR that otherwise would recognize the end of the chromosomes, notably the G-overhand as DNA damage (515–517).

1.3.2.2 ROS

Energy is a necessity in every system, from classical mechanics to quantum mechanics passing by chemistry and life science. The cell is not an exception, and it also requires energy to function. This energy is produced in the mitochondria, that is why Dr. Willis called them the "powerhouse of the cell" in 1992 (518). The conversion of this energy is done via the Krebs cycle, and the energy is presented in the form of Adenosine triphosphate (ATP) (519). Unfortunately, as a byproduct of this cycle, many ROS are produced. The most common examples of these reactive chemicals are the superoxide ($*O^-_2$) or hydrogen peroxide (H_2O_2). The hydrogen peroxide can be partially reduced to hydroxide ion (HO^-) and hydroxyl radical (*OH) (520). Since these molecules have an unpaired electron in their atomic orbital, rendering them unstable, they can attack other stable molecules to accept or donate electrons (521). This process creates a vicious circle where the attacked molecules can, in their turn, become reactive and attack other molecules. It is worth noting that while ROS can attack the cellular components like proteins, lipids, and the DNA, they can also attack the mitochondrial components worsening the vicious cycle. These ROS can be counteracted with a plethora of antioxidants like the manganese superoxide dismutase (MnSOD) (522), glutathione peroxidase (GPx) (523), Glutathione reductase (GR) (524), thioredoxin 2 (525,526), and catalase (CAT) (527). These endogenous antioxidants can be helped by exogenous ones like vitamin C, vitamin E, carotenoids, alkaloids, and polyphenols (528).

During aging, the balance between ROS and antioxidants is tipped towards ROS, causing more and more damage to the mitochondria (529), and the cell, and thus the organism as a whole (530,531). In order to prove the importance of these ROS on aging a ROS-scavenging mitochondrial-targeted catalase was overexpressed in mice, these mice lived longer and healthier and had protection against age-related pathologies like cardiomyopathy (532–534). Since neurons have a high consumption of energy, and since the primary substrate of that energy in the brain is glucose, it is not surprising that the brain is profoundly affected by this imbalance during aging, leading to neurodegeneration with age (535–538). Indeed, we can find many mitochondrial DNA mutations in the brain of AD patients (539), leading to more imbalance and more damage incurred from these ROS.

1.3.2.3 Deregulation of the *Ink4A/Arf* locus

The *Ink4A/Arf* locus is found on chromosome 9 in the human genome. This locus code for two important proteins that control the cell cycle: the $P16^{Ink4A}$ and $P19^{Arf}$ (430). We discussed earlier the importance of these proteins and their role in the cell cycle progression. This locus loses its transcription silencing with age (540). While this derepression can be linked to BMI1, other factors can also influence it like the increased expression of protein C-ets-1 (ETS1), protein C-ets-2 (ETS2), or myeloid/lymphoid leukemia protein 1 (MLL1) (541)

1.3.2.4 Heterochromatin loss

We have discussed in previous chapters the importance of chromatin compaction. In this section, we will explore the link between this compaction and aging. The connection between aging and heterochromatin loss or chromatin reorganization stems from experiments done in the *Saccharomyces cerevisiae* that shows the critical role that heterochromatin plays on the lifespan of this budding yeast (542). This correlation between chromatin loss and age is also observed in *C. elegans*, along with the disruption of nuclear architecture due to the loss of peripheral heterochromatin (543). In order to confirm this association, scientists investigated more complex organisms, the *Drosophila*, an organism that is a model for aging. In *Drosophila*, heterochromatin loss was associated with age, along with an aberrant transcription of genes and transposable elements (544–546). Indeed, the phenotypes observed in *Drosophila* were also found in mammals with old cardiac tissue overexpressing the major satellite repeats due to the loss of chromatin repression (547). Aged human fibroblasts also exhibited a loss of chromatin compaction and a defect in nuclear architecture that was lamin dependent, just like the yeast model (548). Due to the varieties of these modifications in humans, the loss of heterochromatin was put under the umbrella of epigenetic drift, this term was established from an extensive study comparing the epigenome of monozygotic twins. The authors of this study describe how the epigenome of these twins is similar at a young age and become more and more distinct with age (549). Longevity studies of monozygotic twins (550,551) have concluded that around 70% of the difference in life span is due to this epigenetic drift (552). These observations and phenotypes of nuclear disorganization also explain the progeroid syndromes like Hutchinson-Gilford progeria syndrome (HGPS). These patients have a mutation in the *LMNA* gene that encodes the LAMIN A/C (553,554)

and manifest premature aging with osteolysis, atherosclerosis, aged facial features, and myocardial infractions reducing their life expectancy to 12 years old (555). This epigenetic drift can be attributed to many external factors like lifestyle, nutrition, and physical activities (556,557). At the same time, the chromatin loss linked to age can be explained by intrinsic factors such as reduced level of histones proteins with age (558–560), this histone shortage leads to the replacement of core histones by histone variants (561–564).

1.3.2.5 DNA damage accumulation

While the cell can replace nearly all its components from the proteins to the lipids, it does not have a mechanism to replace its DNA. That is why DNA repair machinery is vital in order to combat aging. It is estimated that spontaneous decay of the DNA that is the typical everyday damage not linked to a specific damaging agent is very considerable in the cell (565). This spontaneous decay alone can be estimated to produce 100 000 lesions in every cell, every day, all of this not counting the damage from exogenous sources (566). These damages should be repaired by very efficient machinery to maintain the integrity of the genome. While this machinery in human cells is very developed, a meta-analysis study showed a high correlation between age and DNA damage in humans (567), most probably owing to the accumulation of everyday remainders of damage not repaired every day. This machinery is so important to combat aging that most of the progeria syndrome, except HGPS discussed earlier, have a mutation affecting a prominent protein of DDR. From these diseases we can name Werner syndrome (WS) caused by a mutation in a helicase called Werner (*WRN*), Fanconi anaemia (FA) also caused by a mutation of helicases especially from the Fanconi anemia helicases family like the Fanconi anemia group J protein (*FANCJ*) (568), this helicase is also responsible for another progeroid syndrome called bloom's syndrome (BS) (569), and we have XP that have a mutation in the genes coding for a protein in the nucleotide excision repair (NER) machinery (570,571).

For the sake of clarity, each of these factors was presented as separate entities, but the fact remains that all these factors can influence each other acting as echo chambers, for example the ROS damage is more important if the chromatin is decompacted (572). The complex interaction between these factors leads ultimately to more DNA damage and accelerated aging.

Neurons are very slowly replaced in the brain, if ever, compared to the other cells of the body. That is why the brain is on the frontline of the battle with age, and many neurodegenerative diseases have aging as the most important risk factor.

1.3.3 Alzheimer disease

Dementia is a syndrome affecting the brain and reducing the memory capacity, along with thinking and the ability to perform mundane routine everyday tasks. While many forms of dementia exist, we are going to focus on Alzheimer's disease since the World Health Organization (WHO) estimated that it does represent 60 to 70% of all the cases of dementia in their last report about dementia (573).

In 1901 Dr. Alois Alzheimer diagnosed his patient Auguste Deter with a new disease now called Alzheimer disease in his honor (574). The 50-year-old patient was paranoid and had sleep troubles. Furthermore, the patient suffered from a fast pace of deteriorating memory and increased confusion. Upon the death of the patient, Dr. Alzheimer was able to study the brain following the autopsy further. The postmortem analysis revealed histological anomalies later identified as amyloid plaques and tau neurofibrillary tangles along with massive general atrophy and degeneration of the cortex (575). At the time, the discoveries of Dr. Alzheimer were received with a lack of interest, but what was an isolated case is now considered as the next big challenge of societies. In fact, from what was an isolated case of Auguste Deter in 1901, the diagnosis of this disease kept on growing, and nowadays it is estimated that 5.7 million Americans live with this disease with a doubling time of the number of cases expected to be of 5.5 years (576). This number is estimated to keep growing in the next years to hit 14 million in 2050, according to the American Association of Alzheimer (577).

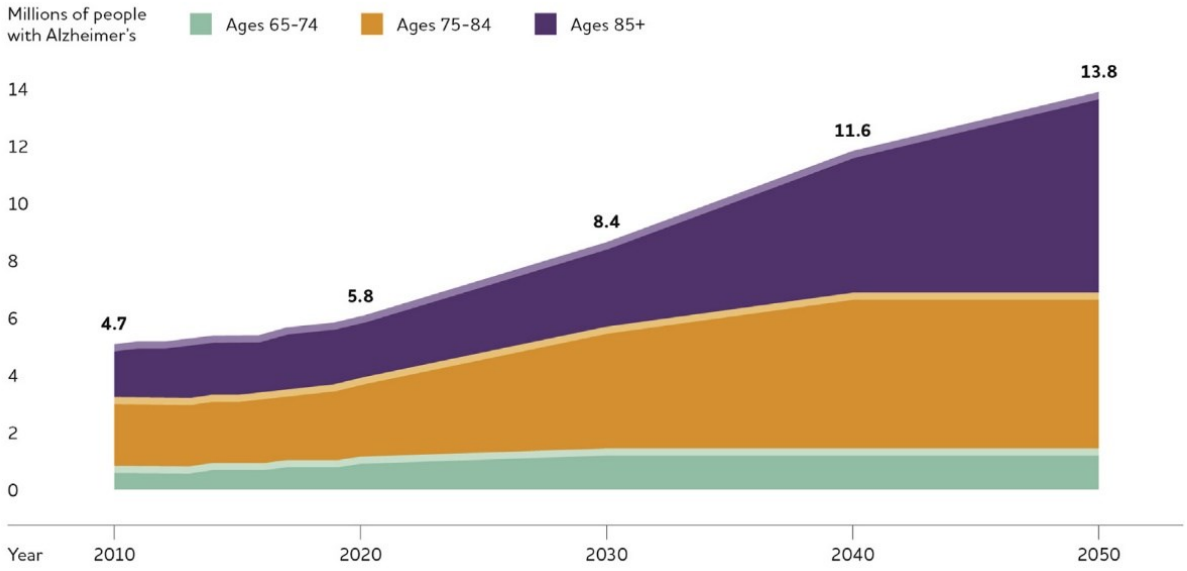


Figure 14. Projection of patient's number in the coming years (577).

This increase in the number of cases comes with a high burden. In fact, it is estimated that in 2017 the USA paid 277 billion dollars dedicated to the care of people with this disease. This number should also be compounded with another as substantial cost, these patients are often, in 48% of the cases, looked after by family members or significant person, these caregivers are not paid, but the estimated cost of their collective work is expected to be around 232.1 billion dollars in 2017. This is not merely an economic problem, the caregivers and the patients have another high cost to pay, a psychological and emotional cost for obvious reasons: The patients do not like to be a burden on the others and carry a feeling of guilt. For any caregiver, it is not easy to take care of a loved one. However, when combining this fact with the memory loss of the loved one that in some cases, they do not even recognize you anymore, this burden becomes even more substantial. As a matter of fact, caregivers incur tremendous emotional, psychological, and physical hardship. This burden is then translated in an increased in their own medical expenses: the average annual health care payment for a person that is not taking care of an Alzheimer patient is around 13 700$ in the USA, this number is almost quadrupled for a caregiver reaching 48 000$ (577). The overall economic cost of AD in the USA 1 trillion$ in 2018 and will hit 3 trillions $ in 2030 (578). All these costs come on societies that are already struggling with their infrastructure and the number of healthcare professionals that are barely enough nowadays and

a population that is aging with less and less workforce and more senior citizens with the percentage of the population aged 65 years old and older is expected to increase according to the American census (see Fig. 15).

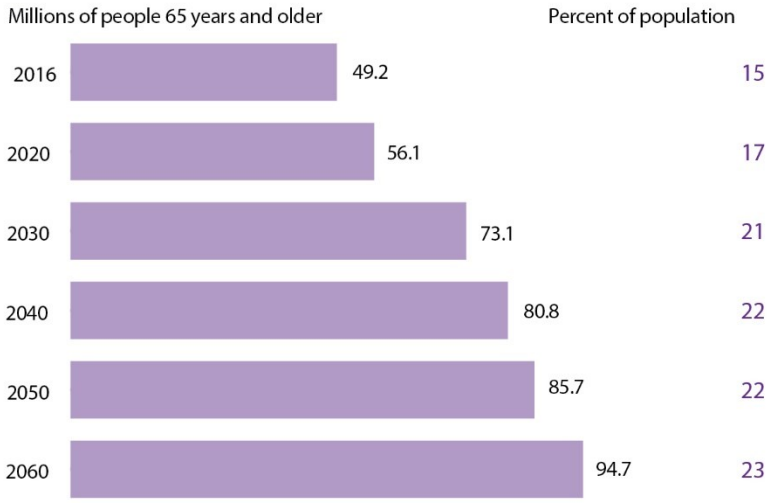


Figure 15. Projection of percentage of the population aged 65 years old and more according to the American census (579)

It is worth noting that this aging of the population is the primary driver of the increase in the number of cases of AD. When we normalize the incidents of the disease with age, we find that the incidence of AD is constant since 1990 (see Fig. 16).

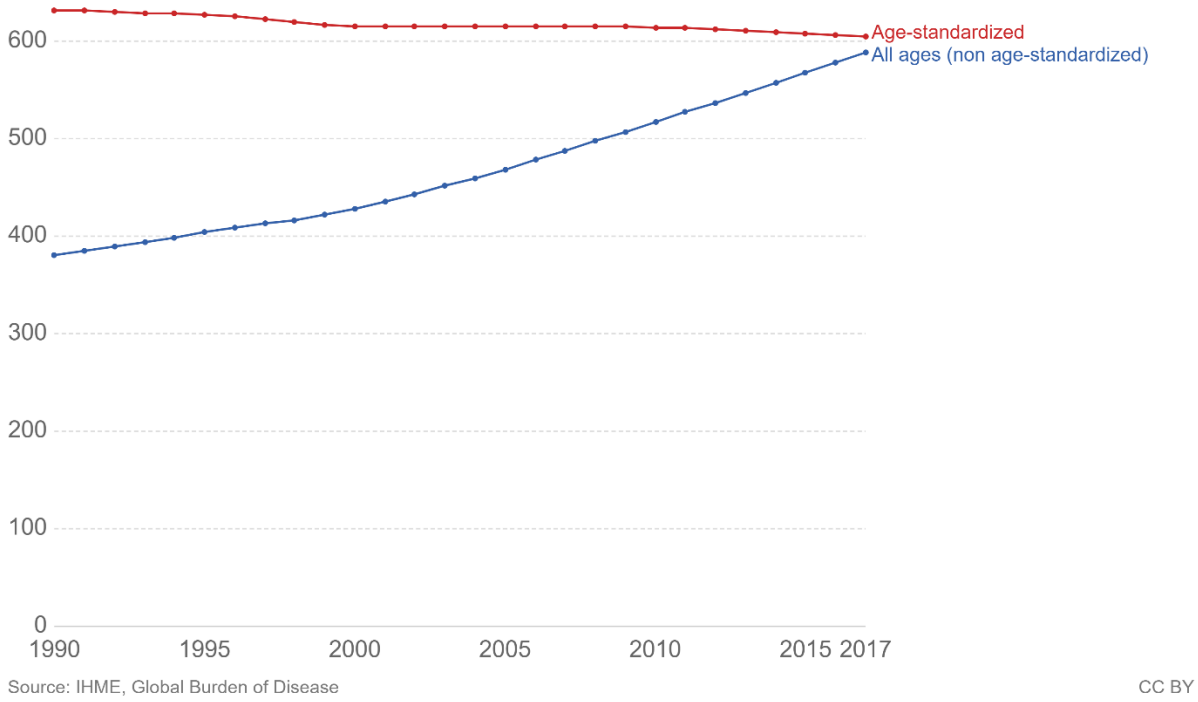


Figure 16. Prevalence of AD over the years (580)

1.3.3.1 AD subcategories

AD is mainly divided into two categories that differ from an etiology point of view and the age of onset point of view. In this chapter, we are going to describe these two distinct diseases.

1.3.3.1.1 Familial AD

Familial AD (fAD), is a hereditary autosomal dominant disease resulting from a mutation of one of three different genes: amyloid precursor protein (*APP*) (581), presenilin 1 (*PSEN1*) (582,583), and presenilin 2 (*PSEN2*) (584,585). Being a disease linked to a specific mutation, the etiology of this disease is well understood. Unfortunately, this category of AD represents less than 5% of all AD cases (586). A significant distinction of this form of AD is the age of onset, having a genetic cause is manifests relatively early for neurodegenerative disease. In this case, fAD onset occurs before the age of 65, that is why this form is also called early-onset Alzheimer disease (EOAD) (587,588). Since Auguste Deter was diagnosed with AD at the age of 51, we now know that she had the EOAD form.

1.3.3.1.2 Sporadic AD

Contrary to the fAD, sAD is not hereditary. The underlying cause of sAD still needs to be elucidated since no mutation was identified to be the cause of this disease. While no specific mutations were identified as disease-causing, Genome-Wide Association Studies (GWAS) have identified many genetic variants to be a risk factor, meaning that the bearer of these variants represents a higher risk of developing sAD than the non-bearer. The most known allele variant identified by GWAS is the *APOEe4*, it is the allele variant that has the most significant effect on the onset of sAD. In fact, recent studies have demonstrated that human neurons derived from iPSC that carried the *APOEe4* variant secreted Aβ40 and Aβ42 peptides, hallmarks of AD, twice as much as the neurons that had the *APOE3* variant (589). Correcting the conformation of APOEe4 using a small molecule protected these neurons and reduced the pathological hallmarks proving that APOEe4 can be a therapeutical target to prevent AD (589). APOEe4 can also fragilize the Blood-brain barrier by activating the CypA-MMP9 pathway leading to a breakdown of this vital barrier that isolates the brain from the rest of the circulatory blood (590). The breakdown of this barrier can lead to the entry of neurotoxic products from the circulatory blood and can increase the inflammation in the brain (591,592), contributing to the cognitive decline seen in AD (590,591). In fact, the breakdown of this barrier is one of the earlier biomarkers of AD (593,594). Other risk factors with smaller effect size are the *TM2D3* gene coding for the Beta-Amyloid-Binding Protein-Like Protein 2 (595), the *COBL* gene coding for the Cordon-bleu protein interacting with actin (596), the *SLC10A2* gene (596), the *TREM2* gene (597), and many other variants. GWAS analysis also identified alleles that were protective against sAD, meaning having this allele reduced the risk of developing sAD. *APOEe2* variant (598), and *APP* (A673T) variants (599,600) are the most important protective alleles against sAD. All these variants were reviewed by Freudenberg-Hua et al. (601), a visual representation of the effect of these alleles, and their prevalence in the population can be seen in figure 17 (reproduced from (601)).

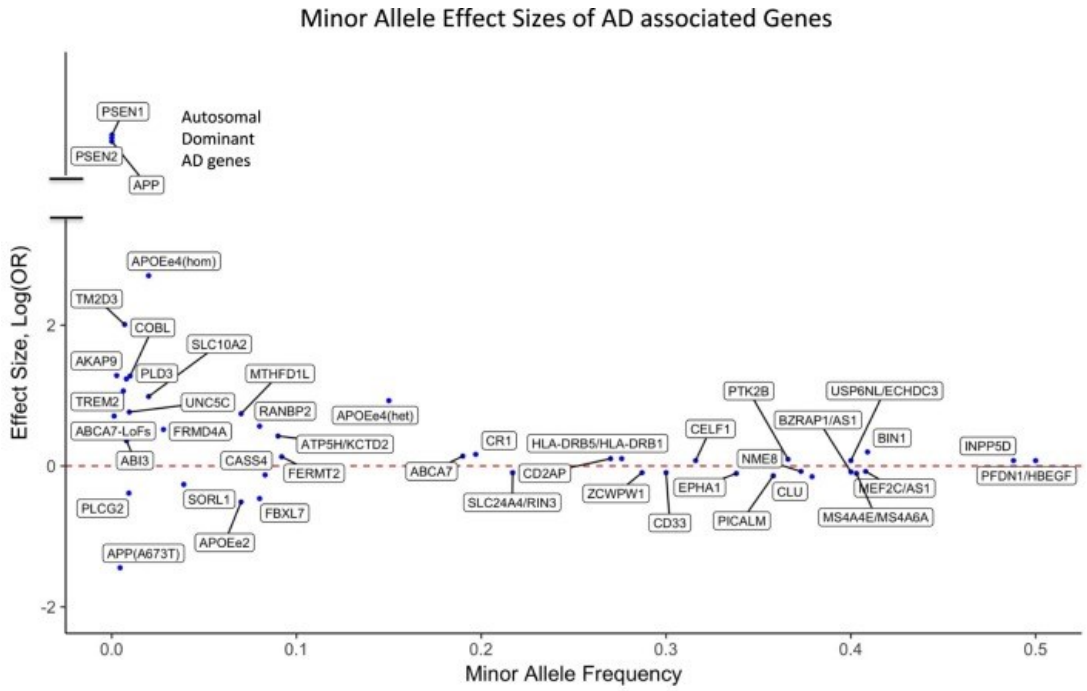


Figure 17. Allele effect on the incidence of sAD with the prevalence of this allele in the population as reproduced from Freudenberg-Hua et al. (601)

Since this form of AD is not hereditary, and we do not have a genetic cause for it, the underlying mechanisms driving this disease are still unknown and need more study, this is unfortunate since 95% of AD cases are sAD. Another significant risk factor of this form is age; in fact, sAD occurs after the age of 65 (587,588); that is why it is also called late-onset Alzheimer's disease (LOAD).

1.3.3.2 Hallmark of AD

While the etiologies of these two forms of AD are different, the physiopathological manifestations are very similar. Below, we are going to reveal the most important hallmarks of AD.

1.3.3.2.1 Brain atrophy

AD brains experience severe atrophy with the progression of the AD disease (602,603). This atrophy is due to neuronal death in AD (604). In the age of Dr. Alzheimer, this atrophy can only be seen in the postmortem autopsy. However, nowadays, these modifications can be visualized with MRI scanning as a tool for diagnostic and a predictive tool for the progression of the disease (605). Examples of brain atrophy can be seen in figure 18. New studies have revealed that this

neuronal death is closely correlated with the progression of dementia and not any other hallmark of AD-like the Ab42 plaques or the Tau tangles (606).

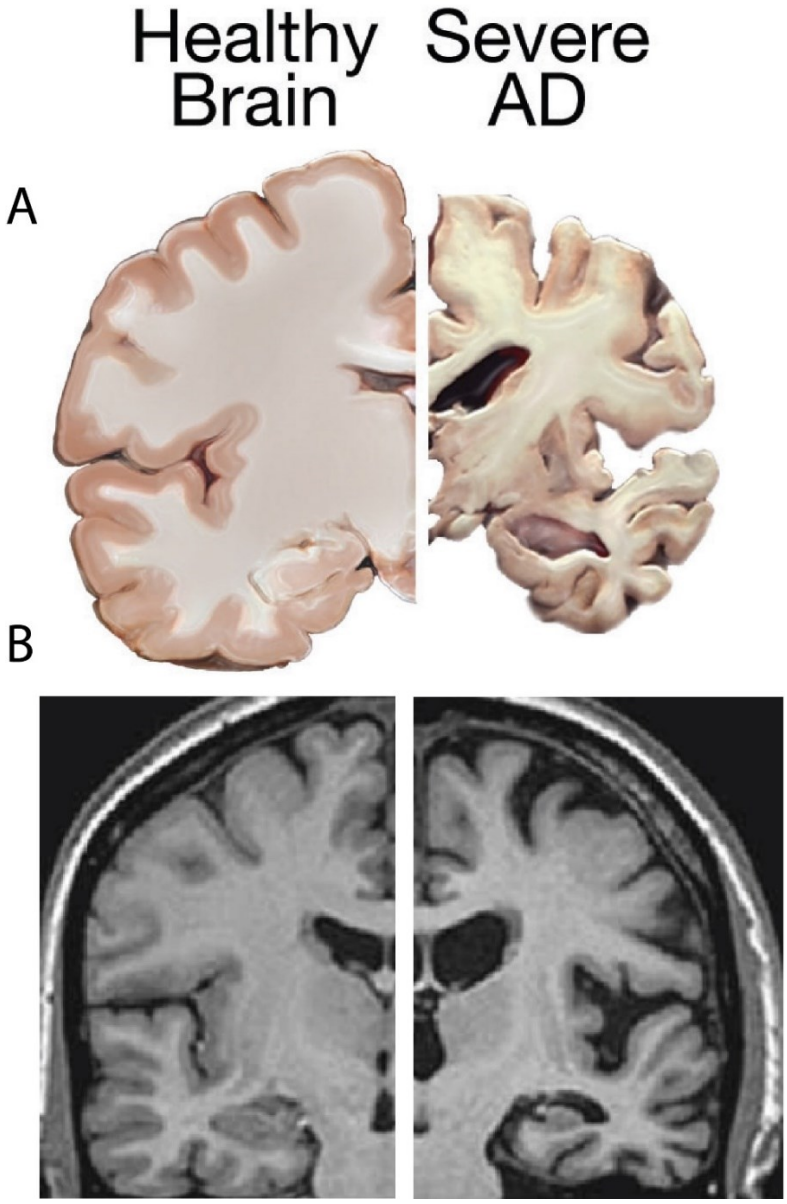


Figure 18. A comparison between a healthy and an Alzheimer's disease brain.

A) represent the picture comparison of the brain postmortem, the AD brain present lots of caveats caused by the atrophy (this picture is in the public domain and was taken from the NIH library (607))

B) represent the MRI results of these a healthy and AD brain. We can notice the caveats even with the MRI (this picture was modified from (608)).

1.3.3.2.2 B-amyloid plaques

APP is a transmembrane protein expressed ubiquitously in the body, but it is known to have a higher concentration at the synapses of the neurons, and most of the known functions of APP are linked to the synapses. It is worth noting that while APP is ubiquitously expressed, APP (695), a variant of APP, is solely expressed in the brain (609). Since a mutation of APP has been identified as one of the causing effects in fAD, most of the studies concentrated on understanding the role of APP in neuron functions. In healthy brains, APP has been shown to play a crucial role in controlling the level of excitation of the synapse's inputs by controlling the transcription levels of keys players in the synaptic transmission pathway (610,611). In another study, the synapses of APP-KO neurons had more synaptic vesicles, and these are the vesicles containing neurotransmitters and released upon the synapses excitation in order to transmit the signal to the other neurons. The increase in the numbers of these vesicles suggests that the APP-KO neurons had more functional synapses than the control neurons, confirming the inhibitory role of APP on the neurons' activity (612).

In the standard physiological processing of APP, or what is called a non-amyloidogenic pathway, APP is first cleaved by an α-secretase (see Fig. 19), this cleavage releases an extracellular peptide called sAPPα and another transmembrane fragment called C83. The cleavage by the α-secretase cut through the Aβ sequence, the sequence responsible for the accumulation, and the aggregation of amyloid plaques. Furthermore, the sAPPα demonstrated neuroprotective properties (613–615). All these observations render the cleavage of APP by an α-secretase beneficial and even neuroprotective. The second step of the non-amyloidogenic pathway is the cleavage of the C83 by the γ-secretases, this cleavage produces an extracellular peptide called p3 and an intracellular peptide called APP- intracellular domain (AICD) (see Fig. 19). We still do not know the exact role of the p3 peptide, but the AICD is a regulator of transcription that can modulate the expression of various proteins involved in the neuronal pathways (616,617). PSEN1 and PSEN2, the two proteins identified in the fAD to be mutated belong to the complex of the γ-secretases, demonstrating the importance of this step in the disease progression.

In a pathological brain, the amyloidogenic pathway is overrepresented. In this pathway, the first cleavage is performed by the β-secretase instead of the α-secretase. This abnormal cleavage

produces an extracellular peptide called sAPPβ and a transmembrane peptide called C99. The C99 fragment contains the intact Aβ sequence that is then released extracellularly when the γ-secretase cleaves the C99 releasing the Aβ fragment and the AICD (Fig. 19). The Aβ fragment is known for its ability to self-assemble, creating soluble oligomers, and the assembly of these oligomers creates the β-pleated sheets observed by Dr. Alzheimer in the brain of Auguste Deter (618,619). The Prevalent theory about AD disease is the "Amyloid theory". This theory states that the imbalance between the amyloidogenic and the non-amyloidogenic pathway results in an increase in Aβ production and thus an increase in Aβ plaque formation, therefore, initiating AD (620,621). Following this theory, many therapeutic human trials were conducted using antibodies targeting Aβ plaques. More than 4 phase III human trials were conducted and failed due to lack of efficacy (622). Although the "Amyloid theory" is the prevalent theory, many experiments challenged it, showing that the presence of plaques is not correlated with the memory impairment and that the removal of these plaques could not stop the progression of AD (623,624). Furthermore, comparison studies revealed that the soluble oligomers of Aβ were the most neurotoxic form of Aβ (625). These new revelations opened the possibility for a new school of thought: in this theory, the plaques are protective against AD because they sequester the toxic oligomers (626,627).

The mechanism behind tilting the balance between the non-amyloidogenic pathway and the amyloidogenic pathway is evident in the fAD since many of the key players are known to be mutated in this form. In sAD, the mechanism behind this switch is more complicated because the key players are not known to have significant mutations.

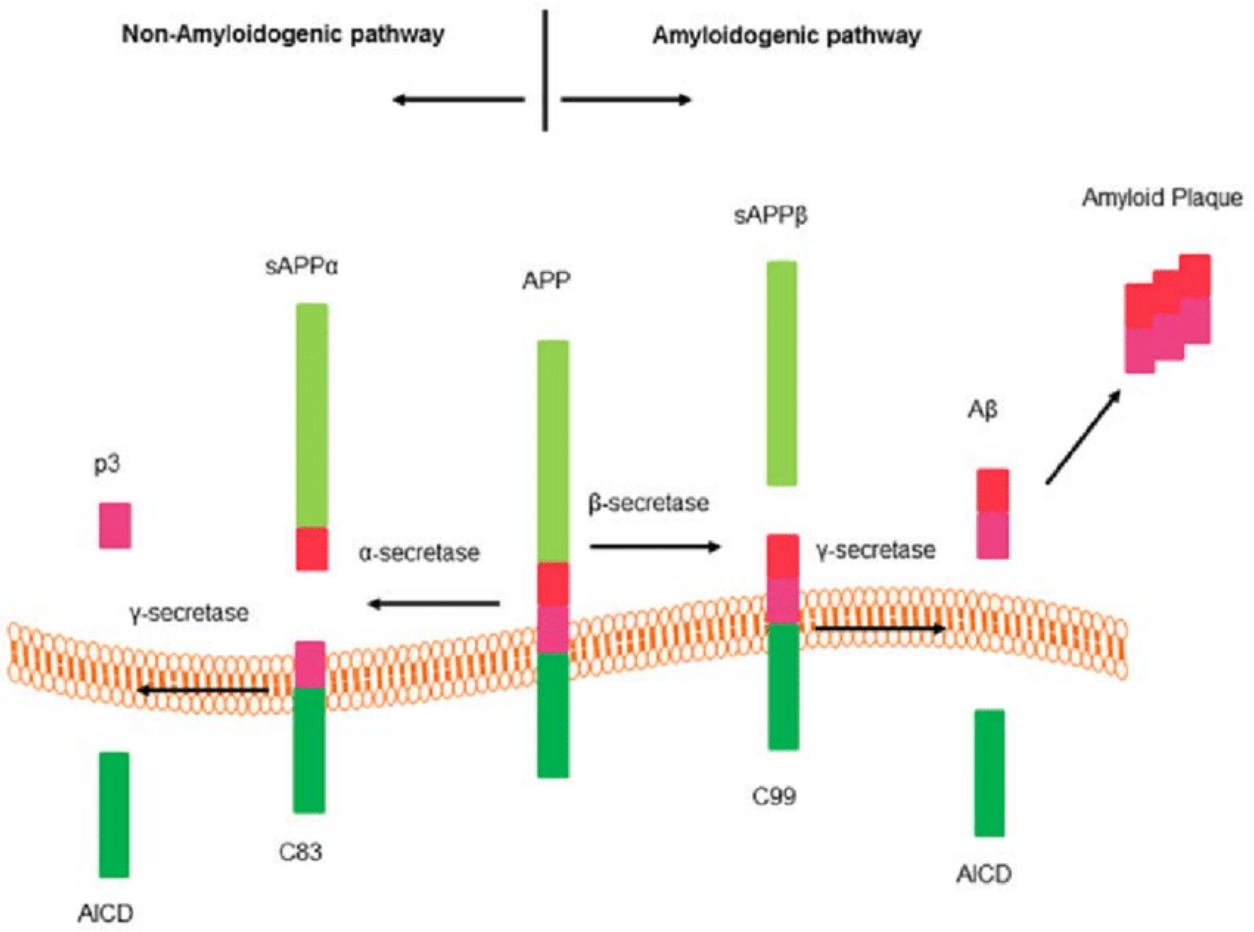


Figure 19. Physiological and pathological pathways of APP processing (628)

1.3.3.2.3 Tau tangles

Tau protein is predominantly expressed in neurons (629). Tau protein is involved in the axoplasmic transport in the neurons (630), this function is especially important in neurons since they are large cells with long axons requiring the transportation of molecules from the nucleus to the effector site: the synapses. Tau binds to the microtubule via its microtubule-binding domain (631); this interaction stabilizes the polymerization of these microtubules (632). Tau can be phosphorylated by different kinases (633). This phosphorylation of Tau reduces the efficiency of tau binding to the microtubule and its effect on stabilizing them (634), the phosphorylation of Tau, and their subsequent disassociation from the microtubules promote Tau self-assembly and oligomerization (635). The oligomerization of Tau produces Tau tangles, also called Neurofibrillary tangles (NFT). NFTs have been identified by Dr. Alzheimer in the brain of Auguste Deter along with

the Aβ plaques as a marker of AD. However, NFTs have been demonstrated in a panoply of neurodegenerative diseases that are grouped under the umbrella of tauopathies.

In AD NFTs are formed before the Aβ plaques in the brain (636), but this accumulation will be restricted to the transentorhinal region of the brain (NFTs stage I and II) (637). NFTs will start invading the limbic regions like the hippocampus (NFTs stage III and IV) and then the neocortex (NFTs stage V and VI) after Aβ plaques deposition (Fig. 20 reproduced from (638)) (638,639). Most importantly, and unlike Aβ plaques, the spatial and temporal deposition pattern of NFTs mimics the neuronal loss (636). Moreover, NFTs deposition, also unlike Aβ plaques, correlate with the cognitive loss progression (640). All these observations taken together make NFTs a better indicator of AD diagnostic and progression than Aβ plaques. It is worth noting that NFTs are considered an indicator of the disease and not the cause of it, in fact, NFT are not toxic (641,642), but the soluble intermediate oligomers of tau are the toxic forms (643,644).

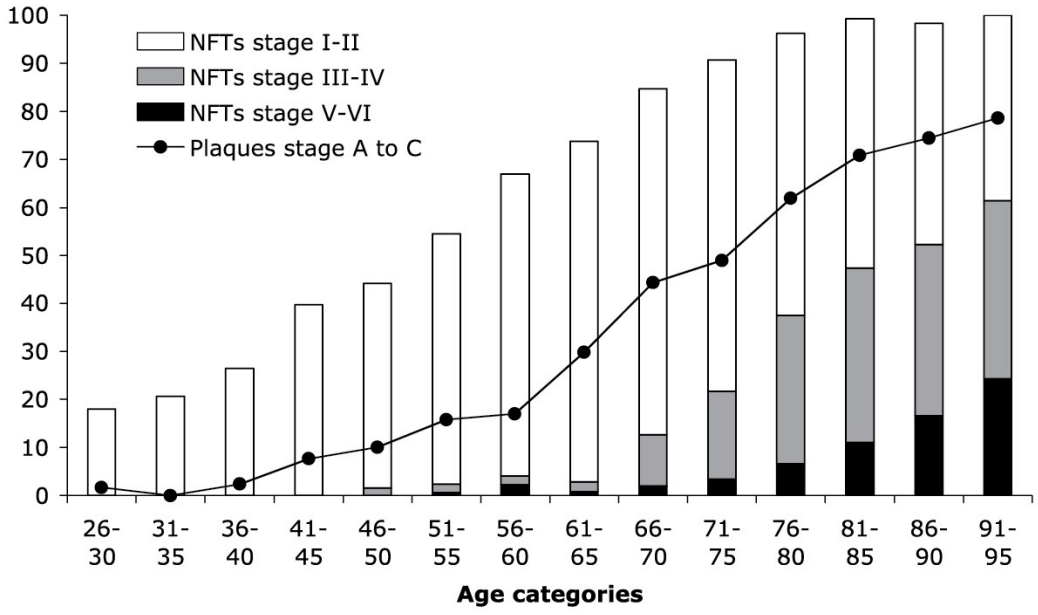


Figure 20. Chronological assessment of the NFTs deposition along with the Aβ plaques (638)

Development of neurofibrillary changes and amyloid deposits in 2,661 non-selected autopsy cases as calculated from (637)

To better understand the trigger initiating AD, researchers studied the influence of Tau on Aβ and vice versa. These studies demonstrated that Tau of its own, without any Aβ, induces the toxicity and the memory loss seen in AD (645–647). Furthermore, the toxicity of Aβ can be mitigated if not rescued by the reduction of Tau levels (648,649). These observations reinforce the theory called “the Tau theory” that states Tau as the primary trigger of AD. Unfortunately, these observations are not clean-cut, and many contradictory pieces of evidence shine a different light, in particular, and experiment that shows that Aβ exacerbates the tau pathology but not vice versa (650) undermining the bases of the “Tau theory”. The reason why Tau is hyperphosphorylated in AD is still to be elucidated.

1.3.4 Sporadic AD and aging

While we do not know how AD is triggered or by what, we do have age as the leading risk factor. Age is, in fact, the common denominator that ties all the scales that we have explained beforehand. Age induces gradual changes in the chromatin structure, these changes cause DNA damage and modified gene expression, all of which can be seen in AD samples. In fact, AD brains have relaxed heterochromatin (447,448). This decondensation of the chromatin leads to the expression at a higher level of the repeat elements of the genome, especially the LINE elements (651,652). The decondensation of the heterochromatin and the subsequent changes in the expression of the repeat element induce, as stated before, DNA damage accumulation, this DNA damage accumulation is well documented in the case of AD (653–656). The accumulation of the DNA damage and the repair machinery of this damage increases the expression of Tau (657,658) phosphorylation of Tau (659). Moreover, to turn full circle, Tau promotes heterochromatin relaxation (660), and p-Tau induces overexpression of the transposable elements (661).

Furthermore, Aβ and tau induce DNA damage (662,663). To conclude, AD is a complex disease in which all the strata of the cell are involved from the microscopic to the macroscopic. Furthermore, that neurodegeneration is a vicious circle in which every stratum feeds the next.

2 Chapter 2 – The Polycomb Repressive Complex 1 Protein BMI1 is Required for Constitutive Heterochromatin Formation and Silencing in Mammalian Somatic Cells

In this article, we studied the role of BMI1 in the maintenance of constitutive heterochromatin in mammalian somatic cells. Before this article, it was thought that BMI1 was excluded from the constitutive heterochromatin and restricted to the facultative one.

While publishing this article we submitted data as “reviewers’ eyes only” since the journal was not publishing supplementary data. The reviewers of our article had access to these figures and assessed the article in light of these figures along with the core manuscript. These data are presented in this work in annex I.

This article demonstrated not just that BMI1 is localized at constitutive heterochromatin, but it plays an essential role in the recruitment of proteins associated with constitutive heterochromatin, specifically to the repetitive loci in order to silence them.

The Polycomb Repressive Complex 1 Protein BMI1 is Required for Constitutive Heterochromatin Formation and Silencing in Mammalian Somatic Cells

Mohamed Abdouh[1£], Roy Hanna[1£], Jida El Hajjar[1], Anthony Flamier[1], and Gilbert Bernier[1]

[1] From the Department of Neurosciences, University of Montreal, and The Stem Cell and Developmental Biology Laboratory, Hôpital Maisonneuve-Rosemont, 5415 Boul. l'Assomption, Montreal, Canada, H1T 2M4

[£] Equal contribution

* Running title: *BMI1 regulates heterochromatin compaction and silencing*

To whom correspondence should be addressed: Gilbert Bernier, Stem Cell and Developmental Biology Laboratory, Hôpital Maisonneuve-Rosemont, 5415 Boul. l'Assomption, Montreal, Canada, H1T 2M4. Tel: 514-252-3400 ext. 4648; Fax: 514-253-7626; E-mail: gbernier.hmr@ssss.gouv.qc.ca

Keywords: Polycomb; BMI1; BRCA1; heterochromatin; histone H2A

Background: BMI1 silences the expression of genes located at the facultative heterochromatin.

Results: BMI1 is abundant at repetitive genomic regions, including the pericentromeric heterochromatin (PCH), where it is required for compaction and silencing.

Conclusion: BMI1 is essential for PCH formation.

Significance: BMI1 function at PCH is important to understand how BMI1 regulates genomic stability.

2.1 Abstract

The Polycomb Repressive Complex 1 (PRC1), containing the core BMI1 and RING1A/B proteins, mono-ubiquitinylates histone H2A (H2Aub), and is associated with silenced developmental genes at facultative heterochromatin. It is, however, assumed that the PRC1 is excluded from constitutive heterochromatin in somatic cells based on work performed on mouse embryonic stem cells and oocytes. We show here that BMI1 is required for constitutive heterochromatin formation and silencing in human and mouse somatic cells. BMI1 was highly enriched at intergenic and pericentric heterochromatin, co-immunoprecipitated with the architectural heterochromatin proteins HP1, DEK1, and ATRx, and was required for their localization. In contrast, BRCA1 localization was BMI1-independent and partially redundant with that of BMI1 for H2Aub deposition, constitutive heterochromatin formation, and silencing. These observations suggest a dynamic and developmentally regulated model of PRC1 occupancy at constitutive heterochromatin, and where BMI1 function in somatic cells is to stabilize the repetitive genome.

2.2 Introduction

Chromosomes are structurally organized in distinct sub-compartments, as determined by the local DNA sequence and chromatin organization. Euchromatin defines “relaxed” chromatin regions containing actively transcribed genes. In contrast, heterochromatin defines “compacted” chromatin regions containing tissue-specific and developmental genes (the facultative heterochromatin) or gene-poor regions (the constitutive heterochromatin) (664). The constitutive heterochromatin is found at the center (centromere) and ends (telomeres) of chromosomes and is mostly constituted of repetitive DNA sequences (664). Numerous (about 10,000) repetitive A/T rich DNA elements of 231 bp are also found in the pericentromeric heterochromatin (PCH) of mouse chromosomes. Because constitutive heterochromatin regions contained repetitive DNA sequences, the maintenance of chromatin compaction is essential to preserve genomic stability (665). During mitosis, repetitive elements can recombine, resulting in non-homologous recombination between different chromosomes or different regions of paired chromosomes and thus chromosomes deletion, translocation, and fusion (666). Repetitive DNA sequences can also be transcribed, resulting in aberrant non-coding RNA. Stabilization of telomeres and centromeres is also essential for chromosome ends capping and kinetochore attachment during mitosis (667). Finally, about 40% of the mammalian genome is constituted of “parasitic” retro-element located in intergenic regions of chromosomes. Active repression of these elements is important to maintain genomic stability since some of these can self-replicate and randomly integrate the genome (668,669).

Nucleosomes are the basic building unit of chromatin and are constituted of a 147 bp of DNA wrapped against a histone octamer containing two molecules of each of the four histones H2A, H2B, H3 and H4 (the nucleosome core particle) (664). The addition of linker histones, such as histone H1, increases the amount of associate DNA by 20 bp to elicit higher levels of chromatin compaction and high order chromatin structure. The chromatin is also attached at multiple points to the nuclear envelope, and the spatial organization of the chromatin in the nucleus is important for the regulation of gene transcription (670,671). Post-translational modifications of histones, such as methylation, acetylation, and ubiquitylation, can modify chromatin compaction and stability. For example, silent or compact chromatin is associated with tri-methylation of histone

H3 at lysine 9 (H3K9me3) or 27 (H3K27me3), while open chromatin is associated with histone H3 tri-methylation at lysine 4 (H3K4me3) or acetylation at lysine 9 (H3K9ac) (672). Acetylation brings in a negative charge, acting to neutralize the positive charge on histones and decreases the interaction of the N termini of histones with the negatively charged phosphate groups of DNA. In pathological conditions, histones hyper-acetylation can result in chromatin and chromosomes de-condensation (673).

A core of proteins is involved in the establishment and maintenance of constitutive heterochromatin. Most of these proteins are conserved in *Drosophila* and were identified as modifiers of position effect variegation (664). In mammals, the buildup of these proteins to heterochromatic DNA follows a relatively well-characterized sequence where zinc finger proteins recognize and bind repetitive DNA sequences. This is followed by enrichment for Histones H1 and H2a/z, accumulation of Hmga1/2, attachment of KAP1/Trim28 (a SUMO E3 ligase), and of the ATP-dependent chromatin remodeler ATR, deacetylation of histones by HDAC2 and tri-methylation of histone H3 at lysine 9 by SUV39h1 and SUV39h2. Association of SUMOylated HP1a and HP1b to a non-coding RNA results in increased binding affinity for H3K9me3, and this is further enhanced and stabilized by the suppressor of variegation protein DEK1 (674,675). The DEK1/HP1/SUV39 complex then propagates the H3K9me3 marks on the chromatin, resulting in heterochromatin formation (676–680). Interestingly, ATRx localizes at both telomeric and PCH, and germline mutations in *ATRx* are associated with the Alpha-Thalassemia with mental Retardation X-linked syndrome (126,681,682). Surprisingly, it was shown that the BRCA1 protein, which possesses histone H2A mono-ubiquitin ligase activity when in complex with BARD1, is also enriched at PCH and required for H2A ubiquitination, heterochromatin compaction and silencing (683–685). This novel BRCA1 function was proposed to explain the severe genomic instability phenotype of BRCA1-deficient cells (684).

Polycomb group (PcG) proteins form large multimeric complexes involved in gene silencing through modifications of chromatin organization (335). They are classically subdivided into two groups, namely Polycomb Repressive Complex 1 (PRC1) and PRC2 (321). Histone modifications induced by the PRC2 complex (which includes EZH2, EED, and SUV12) and the PRC1 complex (which includes BMI1, RING1A, and RING1B/RNF2) allows stable silencing of gene expression in

euchromatin and facultative heterochromatin (344,686,687). Notably, previous recruitment models of PcG proteins through sequential histone modifications have been revised following that PRC1-variants could operate independently- and even upstream of PRC2 (373,374). The PRC2 contains histone H3 tri-methylase activity at lysine 27 (H3K27me3), while the PRC1 contains histone H2A mono-ubiquitin ligase activity at lysine 119 (H2Aub) (344,686,687). A number of observations have implicated these proto-oncogenes in human cancers (688–693). At the opposite, Bmi1-deficient mice display neurological abnormalities, post-natal depletion of stem cells, increased reactive oxygen species, reduced life span, and premature aging phenotypes (286,327,694,695). Likewise, primary human and mouse cells deficient for BMI1 undergo rapid senescence, in part through activation of the tumor suppressor *INK4A* locus (327,695,696). BMI1 was also implicated in DNA damage response and repair, and maintenance of genomic stability (423,697–699).

Although PRC1 proteins have not been directly implicated in constitutive heterochromatin formation or maintenance, at least some line of evidence support this possibility: 1) Immuno-gold localization of BMI1 by electron microscopy in U-2 OS cells revealed high enrichment in electron-dense heterochromatin; 2) BMI1 immuno-localization was found at PCH in transformed human cell lines (356,700–702). However, based on work performed on mouse embryonic stem cells and oocytes, it is generally assumed that PRC1 proteins are excluded from PCH in normal mammalian somatic cells (374,703–705). In contrast with this model, we found that BMI1 is abundant at constitutive heterochromatin in mouse and human somatic cells, and required for heterochromatin formation/maintenance and silencing. In Bmi1-null mice, cortical neurons showed loss of heterochromatin compaction and activation of intergenic retro-elements and satellite repeats. Consistently, Bmi1 co-localized with H3K9me3 and was highly enriched at PCH in mouse neurons. BMI1 was also enriched at constitutive heterochromatin, including PCH, in normal human neural precursors. Furthermore, BMI1 co-purified with architectural heterochromatin proteins and with histone H3K9me3. BMI1 localization and H2Aub deposition at constitutive heterochromatin were EZH2 and H3K27me3-independent. In both transformed and normal primary somatic cells, BMI1 inactivation resulted in the loss of heterochromatin and alteration in the architecture of the nuclear envelope. Notably, BRCA1 localization was unaffected

upon BMI1 deficiency, and both proteins showed partial functional redundancy for H2A ubiquitination, heterochromatin formation, and silencing. These findings reveal an essential function for BMI1 in constitutive heterochromatin formation and silencing in mammalian somatic cells.

2.3 Experimental procedures

2.3.1 Animals

Mice were used in accordance with the Animal Care Committee of the Maisonneuve-Rosemont Hospital Research Center (Approval ID #2009-40; #2009-42; # 2011-23).

2.3.2 Neuronal cultures

Embryonic day 18.5 cortices were dissected in oxygenated HBSS. Following meninges removal, cortices were cut to ~1mm^3 pieces, and incubated at 37ºC for 15 min in 2 ml TrypleEx solution (Invitrogen). Afterward, the enzymatic solution was discarded, and cortex pieces dissociated in HBSS with a 1 ml tip (10 times up and down). After dissociation, cells were plated at 1.5 x 10^5 cells/well on poly-L-lysine-coated 6-well plates or 8-well cultures slides (BD Biosciences). Cells were maintained in a normal medium composed of Neurobasal-A Medium (Invitrogen), Glutamax-I (Gibco), gentamycin (50 µg/ml; Gibco), B27 supplement (Gibco), NGF (50 ng/ml; Invitrogen) and BDNF (0.5 ng/ml; Invitrogen).

2.3.3 Chromatin Immunoprecipitation (ChIP) assay

ChIP was performed using the ChIP Assay kit (Upstate). Cells were homogenized at RT according to the manufacturer's protocol and sonicated on ice for 10 sec at 30% amplitude to shear the chromatin (Branson Digital Sonifier 450, Crystal Electronics, On. Canada). Sonicated materials were immunoprecipitated using 2 µg mouse anti-BMI1, mouse anti-H2AK119ub clone E6C5, mouse anti-RING1B, and mouse anti-HP1 (Millipore), rabbit anti-H3K9me3, and rabbit anti-H3K27me3 (Abcam), rabbit anti-BRCA1 (SantaCruz), and rabbit anti-mouse IgG (Upstate) antibodies. Fragments were then amplified by real-time PCR in triplicates. Human primers sets used were as in (684). ChIP-qPCR data were analyzed according to the Percent Input method. First, the raw Ct of the diluted 1% input fraction is adjusted by subtracting 6.64 cycles (i.e. log2 of the dilution factor 100). Subsequently, the percent input of each IP fraction is calculated according to this equation: $100*2^{(\text{Adjusted Input Ct-Ct(IP)})}$.

2.3.4 Real-time RT-PCR

Mouse cortices or human cells were diced, and RNA was isolated using TRIzol reagent (Invitrogen). Reverse transcription (RT) was performed using 1 µg of total RNA, and the MML-V reverse transcriptase (Invitrogen). Real-time PCR was carried in triplicates using Platinum SYBRGreen Supermix (Invitrogen) and Real-time PCR apparatus (ABI prism 7002).

2.3.5 Micrococcal nuclease and DNase assays

One million (10^6) cells were harvested at the log phase growth and used in either nuclease sensitivity assay. Cells were permeabilized (0.02% l-α-lysolecithin, 150 mM sucrose, 35 mM HEPES, 5 mM KH_2PO_4, 5 mM $MgCl_2$, 0.5 mM $CaCl_2$) on ice for 90 seconds, and then washed in ice-cold PBS. The cell pellet was resuspended in nuclease buffer (150 mM sucrose, 50 mM Tris–HCl (pH 7.5), 50 mM NaCl, 2 mM $CaCl_2$) on ice, and nucleases were added. Digestions were performed at 24°C. Reactions were stopped by adding digestion stop buffer (20 mM Tris.Cl (pH7.4), 0.2 M NaCl, 10 mM EDTA, 2% SDS) and 0.1 mg/ml RNaseA for 30 min at 37°C. DNA was extracted by phenol/chloroform and visualized on 0.8% native agarose gel/ethidium bromide.

2.3.6 Plasmid constructs and viruses

Sequence-specific oligonucleotides stretch shRNA designed to target the BMI-1 ORF (accession #: BC011652): were synthetized. Oligo#1 (nt 1061-1081) 5'-CCTAATACT TTCCAGATTGAT-3', and oligoScramble (nt 573-591) 5'- Ggtacttcattga tgccac-3' were used in this study. These sequences are followed by the loop sequence (TTCAAGAGA) and finally, the reverse complements of the targeting sequences. The double-stranded shRNA sequences were cloned downstream of the H1P promoter of the H1P-UbqC-HygroEGFP plasmid using Age1, SmaI, and XbaI cloning sites. The shRNA-expressing lentiviral plasmids were cotransfected with plasmids pCMVdR8.9 and pHCMV-G into 293FT packaging cells using Lipofectamine (Invitrogen) according to the manufacturer's instructions. Viral containing media were collected, filtered, and concentrated by ultracentrifugation. Viral titers were measured by serial dilution on 293T cells, followed by microscopic analysis 48 hr later. For viral transduction, lentiviral vectors were added to dissociated cells prior to plating. Hygromycin selection (150 µg/ml) was added 48 h later. shBRCA1 constructs (MISSION shRNA) are from Sigma, and siRING1B (FlexiTube siRNA) are from Qiagen.

The RNAi-resistant BMI1-Myc construct (BMI1$^{myc\text{-}R}$) was generated by synthesis (GenScript) and where the nucleotide sequence of the human BMI1 cDNA (5'-CCTAATACT TTCCAGATTGAT-3') was changed to (5-CC**C** AA**C**AC**A**TT**T**CA**A**AT**A**GA**C**-3), thus preserving the original amino-acid sequence of BMI1.

2.3.7 Proteomics

293T cells were transfected with the EFv-CMV-GFP (GFP-293T) or EFv-BMI1-Myc-CMV-GFP (Myc-293T) plasmids. Protein extracts were subjected to immunoprecipitation using an anti-Myc antibody. Immunoprecipitates were resolved by SDS-PAGE, and LC-MS analysis was performed.

2.3.8 Fixation, sectioning, and immunolabeling

Tissues were fixed in 10% buffered formalin and embedded in paraffin according to standard protocols. 5 to 7 µm thick sections were mounted on Super-Frost glass slides (Fisher Scientific) and processed for immunohistochemistry staining. Formalin-fixed paraffin-embedded slices were analyzed by using the Vectastain® ABC kit (Vector) according to the manufacturer instructions. Peroxidase substrate DAB (brown) (Sigma). Observations were made under a fluorescence microscope (Leica DMRE, Leica Microsystems), and images were captured with a digital camera (Retiga EX; QIMAGING; with OpenLab, ver.3.1.1 software; Open-Lab, Canada). Antibodies used in this study were mouse anti-BMI1 and anti-HP1 (Millipore), rabbit anti-H3K9Ac, and anti-H3K9me3 (Abcam). Secondary antibodies used were FITC-conjugated donkey anti-mouse and rhodamine-conjugated donkey anti-rabbit (Chemicon).

2.3.9 Immunoprecipitation and Western blot

For BMI1/Myc immunoprecipitation experiments, 293T cells were transfected with EFv-/CMV-GFP or EFv-BMI1^{Myc}/CMV-GFP plasmids using Lipofectamine according to the manufacturer's instructions. Whole-cell extracts were collected in immunoprecipitation (IP) buffer [100 mM Tris-HCl, pH7.5; 150 mM NaCl; 0.1% Tween 20; protease inhibitors Complete (Roche Applied Science)]. Following the determination of protein concentration, lysates were subjected to immunoaffinity purification. Briefly, protein extracts (4 mg) were incubated with continuous rotation for 3 h at 4°C with 50 ul of affinity matrix carrying mouse monoclonal anti-c-Myc IgG (clone 9E10; Covance).

The matrix was washed four times with the wash buffer (50 mM Tris-HCl, pH 7.5; 150mM NaCl; 0.1%Tween 20; protease inhibitors Complete). The bound proteins were eluted by treating the beads twice with 1 bead volume (50 ul) of c-Myc peptide solution (Covance) (400 ug/ml in 20 mM Tris-HCl, pH7.5, 150 mM NaCl) for 15 min and used in Western blot experiments. Detection and identification of immunoprecipitated proteins were performed by Western blot and LC-MS/MS (liquid chromatography-tandem mass spectrometry). For LC-MS/MS analysis, proteins were revealed in silver-stained gels according to standard protocol. Protein band cutting, trypsin-based in-gel protein digestions, and subsequent LCMS/MS procedures were performed in the Innovation Centre at Genome Quebec.

2.3.10 Primer sequences

Gene	Forward	Reverse
Hprt	5'-ACTGTAATGATCAGTCAACGGG-3'	5-GGCCTGTATCCAACACTTGG-3'
Bmi1	5'-GGAGACCAGCAAGTATTGTCCTATTTG-3'	5'-CTTACGATGCCCAGCAGCAATG-3'
P16	5'-CAACGCCCCGAACTCTTTC-3'	5-GCAGAAGAGCTGCTACGTGAAC3'
Line	5'-TGGCTTGTGCTGTAAGATCG-3'	5'-TCTGTTGGTGGTCTTTTTGTC-3'
Sine	5'-GAGCACACCCATGCACATAC-3'	5'-AAAGGCATGCACCTCTACCACC-3'
Min.sat	5'-TTGGAAACGGGATTTGTAGA-3'	5'-CGGTTTCCAACATATGTGTTTT3'
Maj. sat	5'-GGCGAGAAAACTGAAAATCACG-3'	5'-CTTGCCATATTCCACGTCCT-3'
IAP1	5'-CGCTCCGGTAGAATACTTAC-3'	5'-TGCCATGCCGGCGAGCCTGT-3'

Table 7 . Mouse RT-qPCR primers

Gene	Forward	Reverse
Line	5'-TGGCTTGTGCTGTAAGATCG-3'	5'-TCTGTTGGTGGTCTTTTTGTC-3'
Sine	5'-GAGCACACCCATGCACATAC-3'	5'-AAAGGCATGCACCTCTACCACC3'
Min. sat	5'-TTGGAAACGGGATTTGTAGA-3'	5'-CGGTTTCCAACATATGTGTTTT-3'
Maj. sat	5'-GGCGAGAAAACTGAAAATCACG-3'	5'-CTTGCCATATTCCACGTCCT-3'
IAP1	5'-CGCTCCGGTAGAATACTTAC-3'	5'-TGCCATGCCGGCGAGCCTGT-3'
HoxA7.1	5'-GTGGGCAAAGAGTGGATTTC-3'	5'-CCCCGACAACCTCATACCTA-3'
Globin	5'-CAGTGAGTGGCACAGCATCC-3'	5'-CAGTCAGGTGCACCATGATGT-3'

Table 8 . Mouse ChIP-qPCR primers

Gene	Forward	Reverse
GAPDH	5’-TCACCAGGGCTGCTTTTAAC-3’	5’-ATCCACAGTCTTCTGGGTGG-3’
BMI1	5’-AATCCCCACCTGATGTGTGT-3’	5’-GCTGGTCTCCAGGTAACGAA-3’
P16	5’-GGGTTTTCGTGGTTCACATC-3’	5’-CTGCCCATCATCATGACCT-3’
Alu	5’-CCTCAATCTCGCTCTCGCTC-3’	5’-CTCTAAGGCTGCTCAATGTCA3’
McBox	5’-AGGGAATGTCTTCCCATAAAAACT-3’	5’-GTCTACCTTTTATTTGAATTCCCG-3’
Sat a	5’-AAGGTCAATGGCAGAAAAGAA-3’	5’-CAACGAAGGCCACAAGATGTC-3’
Sat III	5’-AATCAACCCGAGTGCAATCNGAATGGAATCG-3’	5’-TCCATTCCATTCCTGTACTCGG-3'

Table 9 . Human RT-qPCR primers

Gene	Forward	Reverse
HOXC13.2	5’-AGCAGAGCTCAGTGGGAGAG-3’	5’-AATTTCAGGCCCACCCTTAG-3’
Globin	5’-GGCTGTCATCACTTAGACCTC-3’	5’-GGTTGCTAGTGAACACAGTTG-3’
Alu	5’-CCTCAATCTCGCTCTCGCTC-3’	5’-CTCTAAGGCTGCTCAATGTCA-3’
McBox	5’-AGGGAATGTCTTCCCATAAAAACT-3’	5’-GTCTACCTTTTATTTGAATTCCCG-3’
Sat a	5’-AAGGTCAATGGCAGAAAAGAA-3’	5’-CAACGAAGGCCACAAGATGTC-3’
Sat III	5’-AATCAACCCGAGTGCAATCNGAATGGAATCG-3’	5’-TCCATTCCATTCCTGTACTCGG-3'

Table 10 . Human ChIP-qPCR primers

2.3.11Statistical analysis

Statistical differences were analyzed using Student's *t*-test for unpaired samples. Two way-ANOVA test was used for multiple comparisons with one control group. In all cases, the criterion for significance (*P*-value) was set, as mentioned in the figures.

2.4 Results

2.4.1 Bmi1 is required for constitutive heterochromatin formation and silencing in mouse cortical neurons

We performed transmission electron microscopy on cortical slices from WT and *Bmi1*$^{-/-}$ mice at postnatal day 30 (P30). Notably, electron-dense chromocenters were smaller, and the nuclear envelope was generally irregular in *Bmi1*$^{-/-}$ neurons (Fig. 21A). By immunohistochemistry (IHC) on cortical sections and using antibodies against H3K9me3 and H3K9ac-a mark of open chromatin, we observed reduced H3K9me3 labeling in *Bmi1*$^{-/-}$ neurons together with increased H3K9ac labeling (Fig. 21B). Immuno-reactivity for HP1, KAP1, HDAC1, and ATRx was also reduced in *Bmi1*$^{-/-}$ neurons, suggesting heterochromatin anomalies (Fig. 21B). Quantitative analysis revealed that the number of H3K9me3-positive chromocenters was reduced in *Bmi1*$^{-/-}$ neurons, while neuron's nuclear diameter was increased (Fig. 21C). Because post-natal neurodegeneration may account for the observed chromatin anomalies, we analyzed cortical sections from WT and *Bmi1*$^{-/-}$ embryos at e18.5. We found that H3K9me3 and HP1 staining were reduced in *Bmi1*$^{-/-}$ neurons, while that of H3K9ac was unaffected, suggesting that histone hyper-acetylation is secondary to defective heterochromatinization (Fig. 21D). Likewise, cultured cortical neurons from *Bmi1*$^{-/-}$ embryos showed reduced H3K9me3 labeling when compared to *Bmi1*$^{+/-}$ littermates (Fig. 21E). By immuno-fluorescence (IF) on P30 brain sections, we observed reduced immuno-labeling for Lamin A/C at the center of *Bmi1*$^{-/-}$ neuron's nuclei when compared to WT, suggesting anomalies in the nuclear envelope (Fig. 21F). Deficiency in constitutive heterochromatin formation can affect repeat-DNA sequences expression. By quantitative RT-PCR (qPCR) analyses, we found increased expression of Major pericentromeric repeats in *Bmi1*$^{-/-}$ mouse cortices and of intergenic LINE elements and Major and Minor pericentromeric repeats in *Bmi1*$^{-/-}$ cultured e18.5 neurons, when compared to WT (Fig. 21G-H). To test if Bmi1 was enriched at PCH in mouse neurons, we performed Chromatin Immuno-Precipitation (ChIP)-qPCR on cultured WT and *Bmi1*$^{-/-}$ e18.5 neurons after 7 days *in vitro.* We found that Bmi1 specifically accumulated at all repeat-DNA sequences and at the Bmi1-target gene *Hoxa7* (Fig. 22A). Notably, while enrichment for H3K27me3 was observed at *Hoxa7* in WT neurons (and slightly reduced in *Bmi1*$^{-/-}$ neurons),

H3K27me3 enrichment at repeat-DNA was negligible in both conditions (Fig. 22A). *Bmi1*-deficiency in mouse neurons also resulted in depletion of RING1B, HP1, H3K9me3, and H2Aub at repeat-DNA sequences, while the accumulation of BRCA1 was unaffected or increased (Fig. 22A). Bmi1 co-localization with H3K9me3 in mouse cortical neurons was confirmed by IF on brain sections at P30 (Fig. 22B). Bmi1 antibody specificity was further validated by IF on cultured e18.5 cortical neurons (Fig. 22C). These results revealed that Bmi1 is required for heterochromatin formation and repeat-DNA silencing in mouse cortical neurons and enriched at PCH together with RING1B and BRCA1.

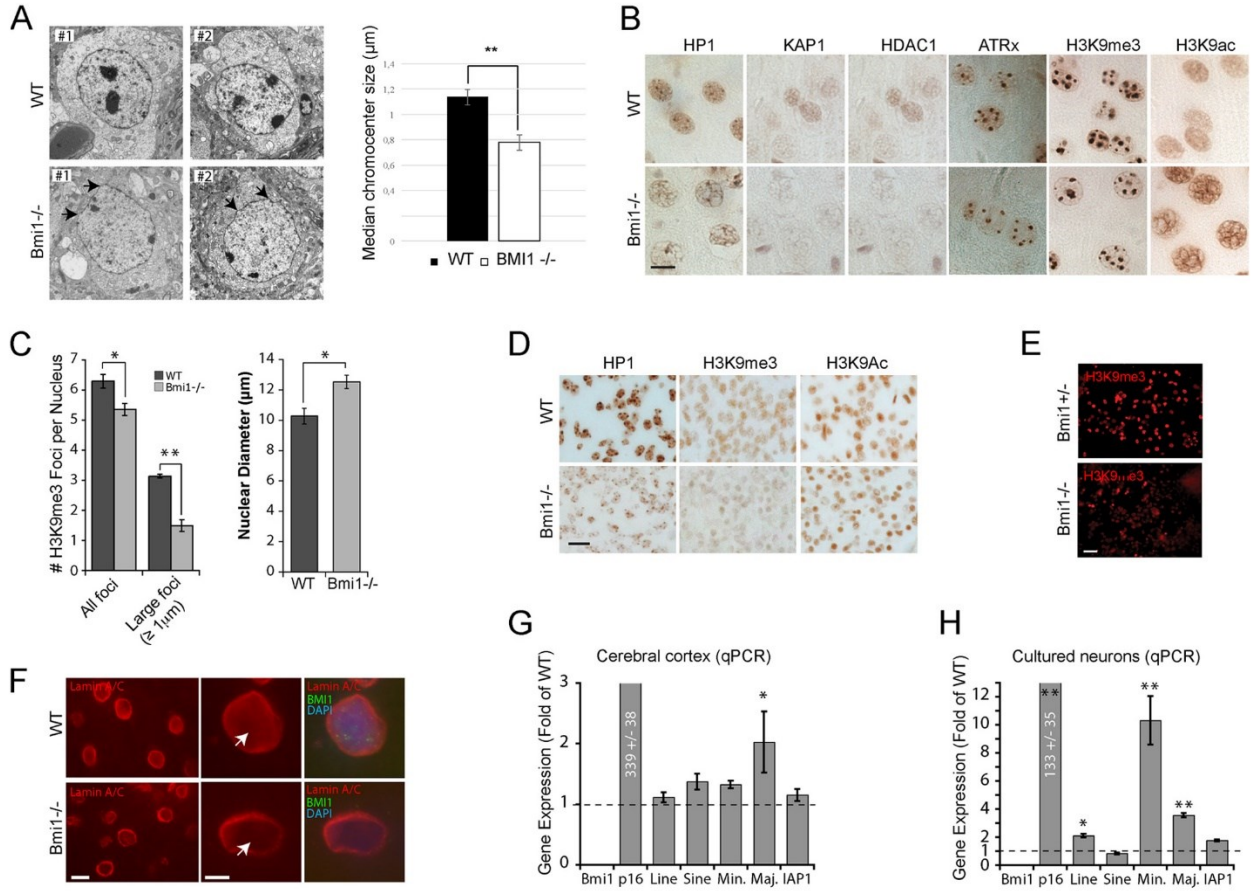


Figure 21. *Bmi1*-deficient mouse cortical neurons present heterochromatin anomalies

(A) Transmission electron microscopy analysis of cortical neurons in P30 WT and *Bmi1*$^{-/-}$ mice. Note the reduction in electron-dense chromocenters and the anomalies in nuclear membrane

architecture in *Bmi1*$^{-/-}$ neurons (arrows). (B) Paraffin-embedded brain sections from P30 WT and *Bmi1*$^{-/-}$ mice were analyzed by immuno-histochemistry. Labeled cells are neurons located in the upper cortical layers of the cerebral cortex. Scale bar, 10 µm. (C) Quantification of the total number of H3K9^{me3}-positive chromocenter and number of large H3K9me3-positive chromocenters. Note that neuron's nuclear diameter is increased in *Bmi1*$^{-/-}$ mice. Where n = 3 brains for each genotype. *P < 0.05, **P < 0.01. (D) Paraffin-embedded brain sections from e18.5 WT and *Bmi1*$^{-/-}$ embryos were analyzed as in (B). (E and F) Cultured embryonic cortical neurons (E) and P30 cortical sections (F) were analyzed by immuno-fluorescence, revealing reduced H3K9me3 and Lamin A/C labeling (arrows) in *Bmi1*$^{-/-}$ neurons. Scale bars: (E) 40 µm, (F) 10 µm, (F') 5 µm. (G) Whole cortices or (H) e18.5 neurons from WT and *Bmi1*$^{-/-}$ mice were analyzed by qPCR for satellite repeats and intergenic retro-elements expression. *P16*^{Ink4a} was used as positive control. Note the up-regulation of minor and major satellite repeats in *Bmi1*$^{-/-}$ neurons. Where n = 3 independent samples for each genotype. *P < 0.05, **P < 0.01.

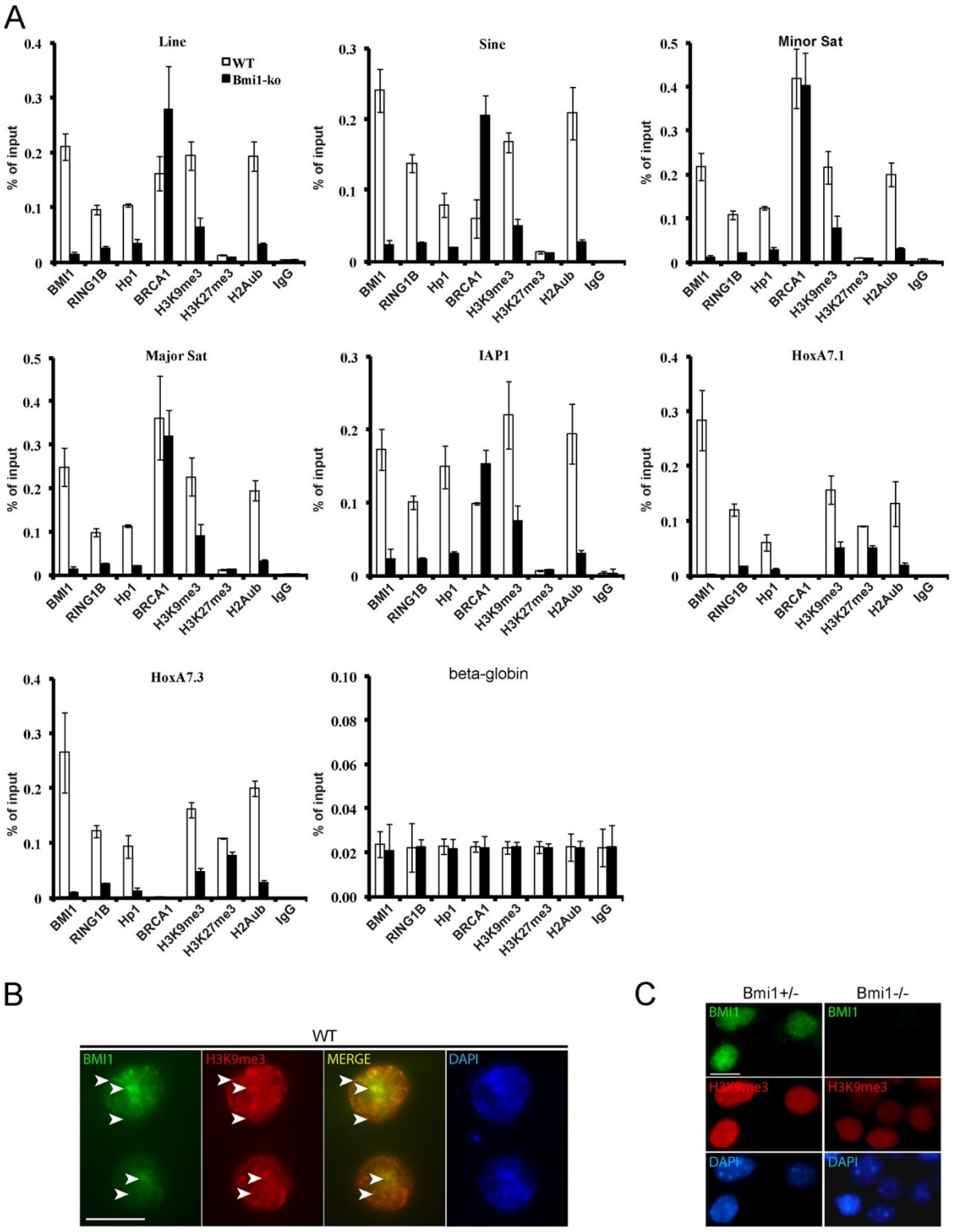


Figure 22. Bmi1 is required for H2Aub deposition and accumulates at repeat-DNA sequences in mouse cortical neurons

(A) WT and *Bmi1*$^{-/-}$ neurons were analyzed by ChIP for proteins enrichment at satellite repeats, intergenic retro-elements and *HoxA7* (positive control). Note the accumulation of Bmi1 and Ring1b at all repeat-DNA sequences, including Major and Minor satellite repeats. While HP1 accumulation and H2Aub and H3K9me3 deposition were reduced in *Bmi1*$^{-/-}$ neurons at all tested loci, BRCA1 accumulation was either unaffected (Minor and Major satellites) or increased (Line, Sine and IAP). Note the near absence of BRCA1 accumulation at *HoxA7.1* and *HoxA7.3* in both WT and *Bmi1*$^{-/-}$ neurons. (B) Immuno-fluorescence analysis showing Bmi1 co-localization with H3K9me3 in WT mouse cortical neurons at P30 (arrowheads). Scale bar, 10 μm. (C) Immuno-fluorescence analysis showing loss of Bmi1 signal and reduced H3K9me3 labeling in cultured e18.5 *Bmi1*$^{-/-}$ mouse cortical neurons when compared to *Bmi1*$^{+/-}$ neurons. Scale bar, 10 μm.

2.4.2 BMI1 is highly enriched at repetitive sequences in human neural precursors

To investigate BMI1 distribution on the chromatin genome-wide, we took advantage of publicly available BMI1 ChIP-Seq raw data on normal human neural precursors (706). Using MACS statistical peak calling, we identified 21,525 BMI1 binding sites. The majority of the peaks (56%) were located at intergenic regions, which are highly enriched for constitutive heterochromatin (Fig. 23A). As expected, we observed BMI1 enrichment at the canonical BMI1 targets *CDKN2A* ($p16^{INK4A}$) and *HOXC* locus (Fig. 23E; top). Among 9,471 gene-associated BMI1 peaks, 714 peaks were also enriched by at least 2-fold for either H3K9me3 or H3K27me3. Notably, 565 were marked for both H3K9me3 and H3K27me3 (Fig. 23B). We further annotated BMI1 peaks surrounding DNA repetitive sequences using repeat masker. Surprisingly, 81% of the total pool of BMI1 was located at repetitive sequences (Fig. 23C) with 985 peaks co-enriched with H3K9me3 and 1,067 peaks co-enriched with H3K27me3 (Fig. 23D). If considering repetitive sequences containing at least one BMI1 peak, LINE, SINE and LTR were the most represented families of repeats. Among satellite repeats, BMI1 was mainly enriched at PCH regions (Fig. 23F-inset). A closer look at PCH regions on chromosome 9 revealed a “pocket-like” deposition of BMI1 peaks surrounding the H3K9me3 deposition (Fig. 23E). A similar pattern of BMI1 peaks distribution was also found at PCH regions in human chromosomes 1-10. Notably, repetitive sequences containing

3 or more BMI1 peaks were largely represented (80%) in centromeric satellite repeats (Fig. 23G). In contrast, although 20% of LINE L1 repeats contained 3 or more BMI1 peaks, no other LINE subfamilies were highly enriched for BMI1 (Fig. 23G). We concluded that in human neural precursors, BMI1 is enriched at constitutive and facultative heterochromatin with prevalence for repetitive sequences.

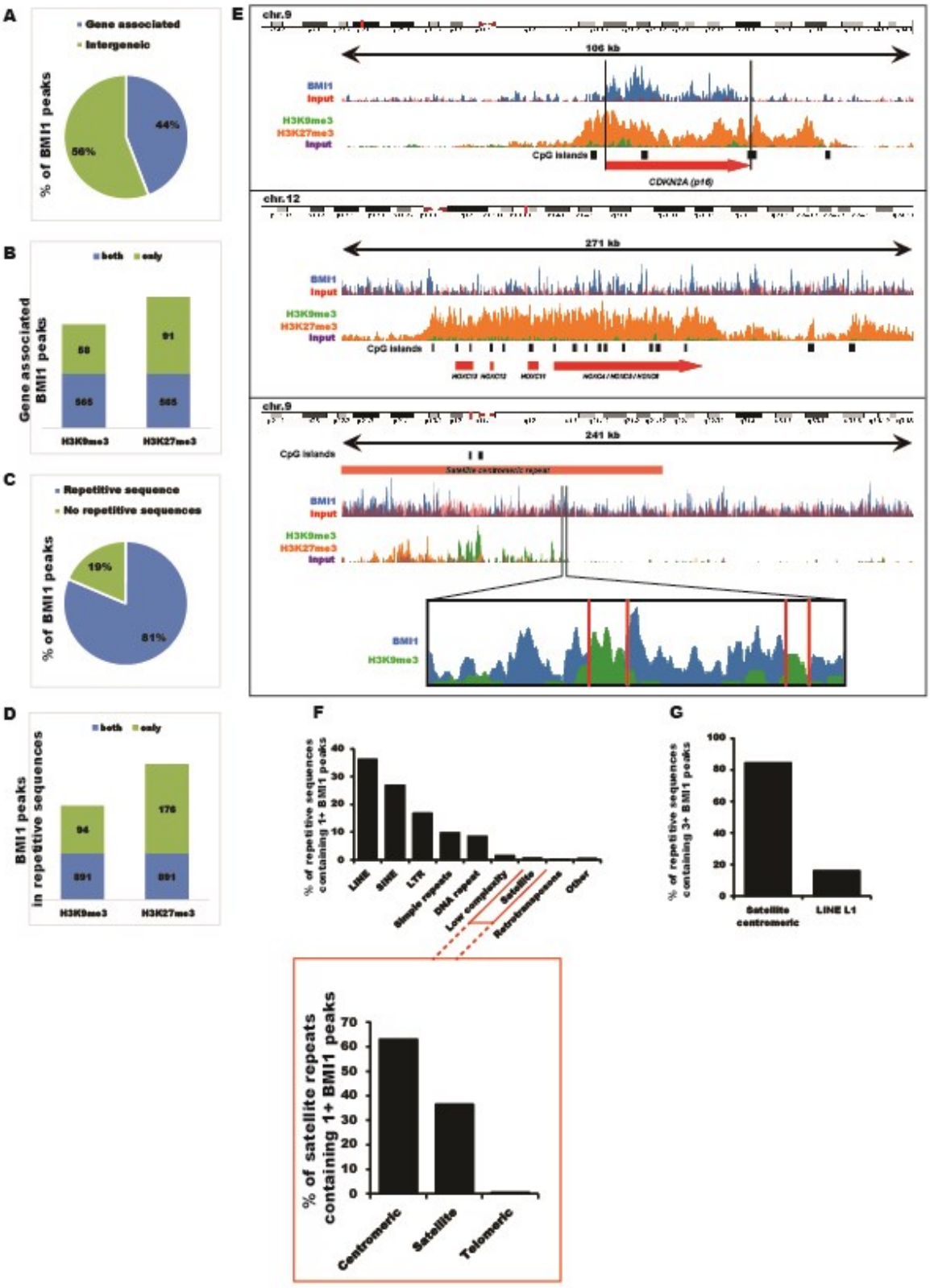


Figure 23. BMI1 is enriched at repetitive sequences in human neural progenitor cells

(A) Proportion of BMI1 peaks associated to a gene (surrounding or upstream 2kb of a gene). Total number of peaks: 21,525 (B) Proportion of gene associated BMI1 peaks co-enriched for H3K9me3 or H3K27me3. (C) Proportion of BMI1 peaks surrounding a repetitive sequence. (D) Proportion of repeat-associated BMI1 peaks co-enriched for H3K9me3 or H3K27me3. (E) Example of canonical BMI1 target genes (*CDKN2A* and *HOXC*) and of the pericentromeric region of human chromosome 9. Red arrowheads indicate BMI1 peaks. Top: physical map on the chromosome. (F) Families of repeat containing at least one BMI1 peak. (G) Families of repeat containing three or more BMI1 peak. BMI1 peaks determined by MACS peak calling; p-value<0.05. At least 2-fold enrichment for H3K9me3 and H3K27me3 were considered.

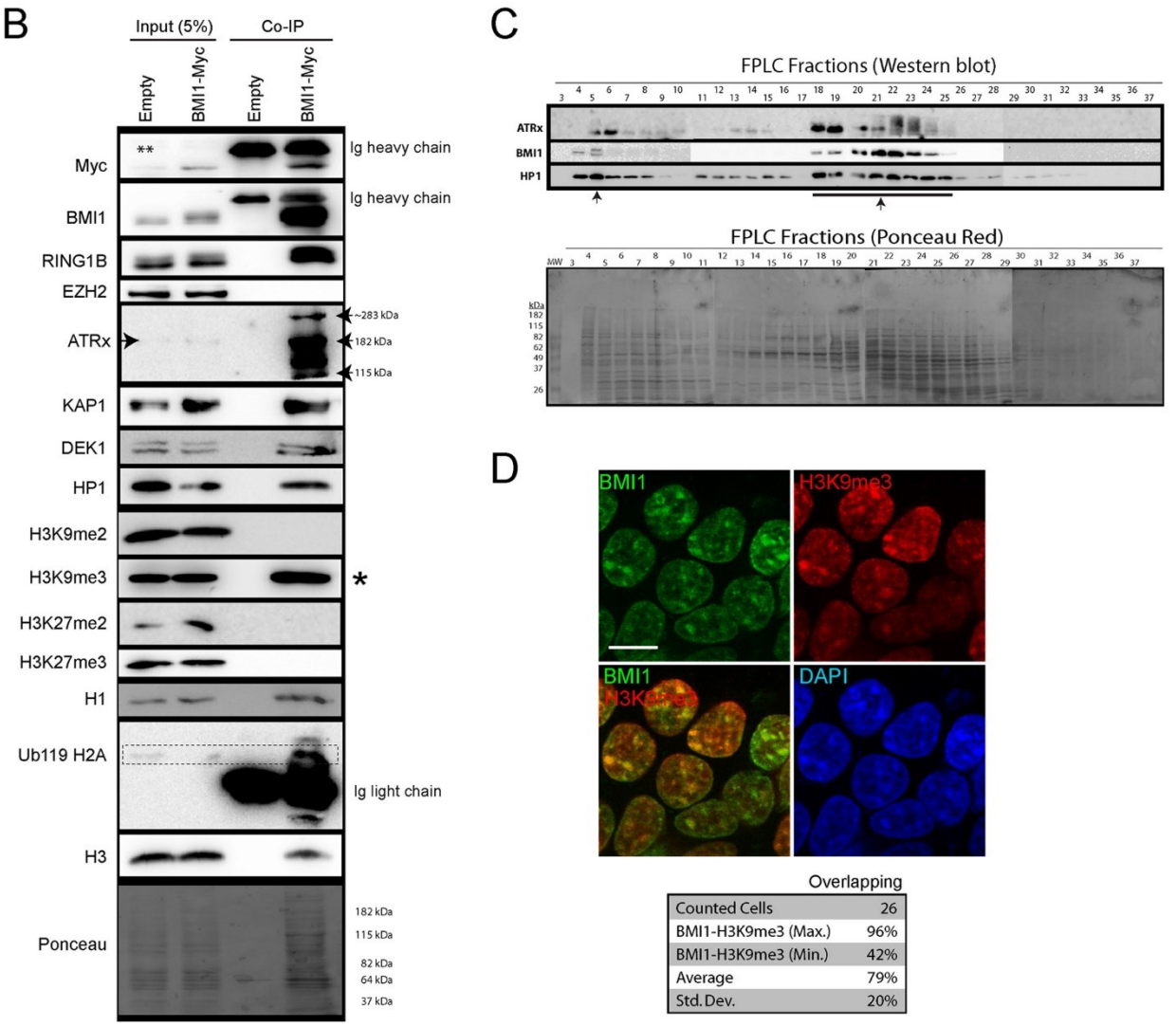
2.4.3 BMI1 co-purifies with architectural heterochromatin proteins

To identify new BMI1 partner proteins, we infected 293T cells with a lentivirus expressing a Myc-tagged BMI1 fusion protein and GFP (EFv-BMI1^{Myc}/CMV-GFP) or a control virus only expressing GFP. After immuno-precipitation (IP) with an anti-Myc antibody, samples were separated on a 1D gel and sequenced by LC-MS (Fig. 24A). We identified several unique peptides in BMI1^{Myc} samples corresponding to proteins involved in heterochromatin organization, including histone H1x, HP1a (also called CBX5), LAMIN A/C and LAMIN B, DEK (also called DEK1) and CENP-V (Fig. 24A) (670,674,707,708). We also identified 2 members of the ISWI-family, BAZ1a (also called ACF1) and BAZ1b, which can promote heterochromatin formation and transcription silencing by generating spaced nucleosome arrays (709). To validate some of these findings, we performed IP experiments on control and BMI1^{Myc} virus-infected cells. As expected, we observed that RING1B, but not EZH2, co-precipitated with BMI1 (Fig. 24B). Co-precipitation of ATRx, KAP1, DEK1 and HP1 with BMI1 was also observed, with a notable enrichment of ATRx when compared to input (Fig. 24B), and where the full length (~280Kda) and truncated (~180Kda) ATRx isoforms were present, together with a lower molecular weight isoform of ~115Kda. Notably, while co-precipitation with histones H3K9me3, H3 (total), H1 and H2Aub was robust, co-precipitation was not observed with histones H3K9me2, H3K27me2 and H3K27me3 (Fig. 24B). We use FPLC to separate protein complexes and found that BMI1 was present in one fraction of very large molecular weight and in several other fractions of lower molecular weight all also containing HP1 and ATRx (Fig. 24C).

By IF studies, we confirmed that BMI1 largely co-localized with H3K9me3 in interphase nuclei (Fig. 24D).

A

Protein Name	NCBI Accession	Molecular Weight(Da)	Identification Probability	%Sequence Coverage	#Unique Peptides
Histone H1x	NP_006017	22,469.90	100%	17.40%	3
CBX5/HP1	CAG33699	22,207.60	99.80%	12.60%	2
LMNB2	NP_116126	67,671.70	100%	15.30%	8
LMNA (Lamin-A/C)	CAI15523	74,053.50	100%	12.30%	8
BAF (Barrier-to-autointegration factor)	NP001137457	10,040.70	100%	42.70%	3
DEK	NP_003463	42,657.90	100%	13.60%	4
BAZ1A/ACF1 (Bromodomain adjacent to zinc finger domain protein 1A)	AAH20636	178,689	99.80%	2.19%	2
NUP93 (nucleoporin)	NP_055484	93,367.10	99.80%	3.17%	2
KPNB1 (karyopherin (importin) beta 1)	NP_002256.2	97,167.40	100%	8.11%	5
KPNA2 (karyopherin alpha 2 (importin alpha 1))	NP_002257	57,844.50	99.80%	7.56%	2
B3KRS5/HDAC2		55,347.50	99.80%	5.94%	2
CENPV (Centromere protein V)	NP_859067	29,928.20	100%	18.20%	3
BAZ1B	AAH65029	170,889	100%	2.36%	3



Overlapping	
Counted Cells	26
BMI1-H3K9me3 (Max.)	96%
BMI1-H3K9me3 (Min.)	42%
Average	79%
Std. Dev.	20%

Figure 24. BMI1 co-purifies with architectural heterochromatin proteins

(A, B) 293T cells were infected with EFv/CMV-GFP or EFv-BMI1^{Myc}/CMV-GFP viruses. Protein extracts were subjected to IP using an anti-Myc antibody, and immunoprecipitates were resolved by SDS-PAGE and analyzed either by LC-MS/MS (A) or Western blot (B). (A) Note the co-purification of BMI1 with several heterochromatin proteins and with Lamins. (B) Note the preferential co-purification of BMI1 with histone H3K9me3 (*) and ATRx. The ** symbol on the panel indicates an artifact coming from partial leakage of the second sample. (C) Native nuclear extracts were size-fractionated by FLPC and analyzed by Western blot (upper panel) and Ponceau Red staining (lower panel). Note BMI1 co-fractionation with ATRx and HP1-containg protein complexes (arrows). (D) 293T cells were labeled with BMI1 and H3K9me3 antibodies, counterstained with DAPI, and analyzed by confocal microscopy. Note the co-localization of BMI1 with H3K9me3-positive chromatin domains. Scale bar, 10 µm. Quantitative confocal analysis was used to measure the proportion of overlapping signals.

2.4.4 BMI1 is required for heterochromatin compaction and silencing

To evaluate BMI1 activity in heterochromatin silencing, we measured gene expression in loss- and gain-of-function experiments. Upon BMI1 deficiency, 293T cells showed reduced proliferation and underwent cell proliferation arrest after 3 passages (Fig. 25A). BMI1 over-expression had however no apparent adverse effect on cell proliferation (not shown). In BMI1 knockdown cells, expression of the canonical BMI1 target gene *p16^{Ink4a}* as well as that of *McBox* and *SATIII* was increased (Fig. 25B). Conversely, BMI1 over-expression resulted in transcriptional repression of *p16^{Ink4a}* and of all tested repeat-DNA sequences (Fig. 25C). By ChIP-qPCR experiments on shScramble and shBMI1-treated cells, we found that BMI1 and RING1B were highly enriched at repeat-DNA sequences and *HOXC13* in control cells (Fig. 25D). In shBMI1 cells, BMI1, RING1B, HP1, H3K9me3 and H2Aub were reduced at all chromatin regions tested (Fig. 25D). In contrast, BRCA1 enrichment at repeat-DNA sequences was independent of BMI1 function. Increased BRCA1 enrichment was even observed at *ALU* sequences upon BMI1 deficiency (Fig. 25D). We tested if RING1B knockdown mimicked the BMI1-deficient phenotype. While RING1B accumulation at constitutive heterochromatin and *HOXC13* was highly reduced in siRING1B-treated cells (clone #4, 80% RING1B knockdown), H2Aub reduction was only detected at *HOXC13* (Fig. 41). RING1B knockdown had no effect on BMI1, HP1 and H3K9me3 on all tested regions. No

significant effect on repeat-DNA sequences expression was observed (Fig. 41), revealing that RING1B knockdown is not sufficient to reproduce the BMI1-deficient heterochromatin phenotype. BMI1 enrichment at repeat-DNA sequences was also EZH2 and H3K27me3-independent (Fig. 42). Nuclease hypersensitivity is a common phenotype of cells deficient in heterochromatin condensation (674,710–712). We used native chromatin extracts isolated from control and shBMI1 293T cells in MNase and DNaseI experiments and found that cells with BMI1 knockdown were hypersensistive to both nucleases (Fig. 25E), thus suggesting globally reduced chromatin compaction.

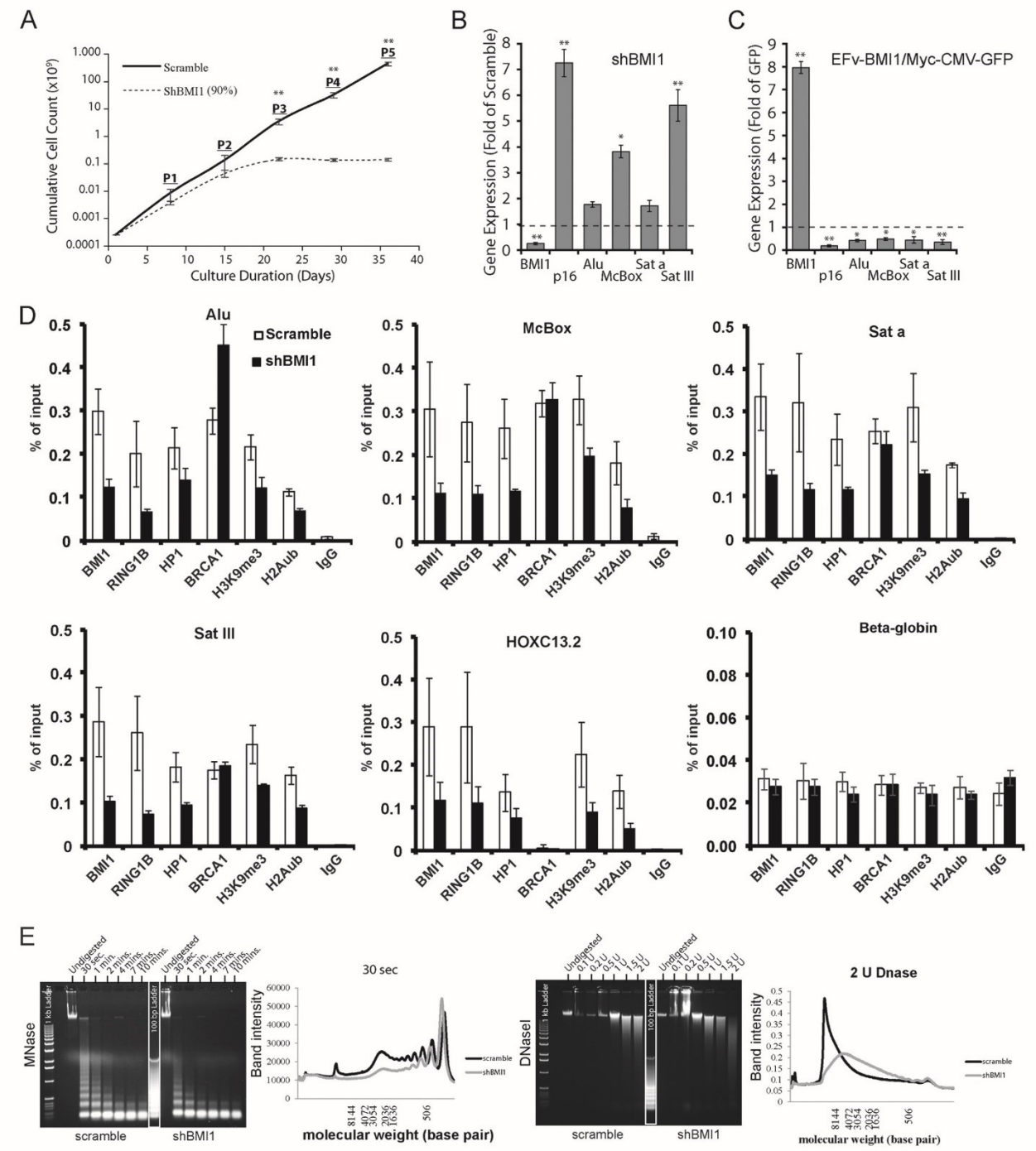


Figure 25. BMI1 is required for heterochromatin compaction and silencing in human cells

(A-C) 293T cells were infected with shScramble or shBMI1 viruses (A, B), or with viruses expressing either GFP or the BMI1-myc fusion protein and GFP (C). (B, C) Gene expression was analyzed by qPCR, and where n = 3 independent cultures. (D) 293T cells knockdown for BMI1 were analyzed by ChIP for proteins enrichment at satellite repeats, intergenic retro-elements and *HoxC13.2*

(positive control). (E) 293T cells were infected with shScramble or shBMI1 viruses and treated or not with MNase (0.4U at 24°C for different time periods) or DNaseI at the indicated concentrations for 20 minutes at 24°C. Note the nuclease hypersensitivity phenotype of BMI1 knockdown cells. *P < 0.05, **P < 0.01.

2.4.5 Severe heterochromatin and nuclear envelope alterations in human cells deficient for BMI1

To further characterize the BMI1-deficient phenotype, we analyzed cells by confocal IF using heterochromatin and nuclear envelope markers. We found severe depletion of the H3K9me3 and H3K27me3 histone marks in BMI1-knockdown cells together with dramatic elevation of the H3K9ac mark (Fig. 26A and C). Likewise, DEK1 and HP1 heterochromatic nuclear foci were lost upon BMI1 knockdown and positive cells for LAMIN A/C were significantly reduced (Fig. 26B-C), suggesting perturbation of the nuclear envelope architecture.

To test whether this correlated with alterations in the subnuclear distribution of heterochromatin proteins, we performed cellular fractionation experiments and where the SDS-soluble fraction is thought to be highly enriched for constitutive heterochromatin proteins (425,674). In controls cells, BMI1 was detected in the 450nM NaCl and SDS fractions (Fig. 26D). ATRx, HP1, BRCA1 and H3K9me3 were also highly enriched in the 100nM-450nM NaCl- and SDS-soluble nuclear fractions. Modest HP1 distribution was also found in the nucleosol fraction (Fig. 26D). In shBMI1 cells, which underwent premature cell proliferation arrest (thus explaining the overall reduced total protein loading), ATRx, HP1 and H3K9me3 were highly reduced in the SDS fraction and displaced in the other fractions. In contrast, the distribution of BRCA1 in chromatin fractions and its overall expression were unaffected upon BMI1 knockdown (Figs. 26D and Fig. 43). To test BMI1 function in primary human cells, human dermal fibroblasts were infected with the lentiviruses and analyzed by IF. We observed that in contrast to control cells where robust H3K9me3 labeling was widespread throughout the interphase nucleus, H3K9me3 labeling in BMI1 knockdown cells was highly reduced and present at the nuclear periphery where it did not co-localized with DAPI (Fig. 26E). Notably, co-localization of H2Aub and H3K9me3 with DAPI in shBMI1-infected cells could be rescue by an RNAi-resistant BMI1-Myc fusion protein ($BMI1^{myc\text{-}R}$),

thus excluding possible off-target effects (Fig. 26F). Bubbling of the nuclear envelope and loss of DEK1 nuclear labeling were also observed (Fig. 26E), revealing BMI1 requirement for constitutive heterochromatin maintenance in both transformed and primary human cells.

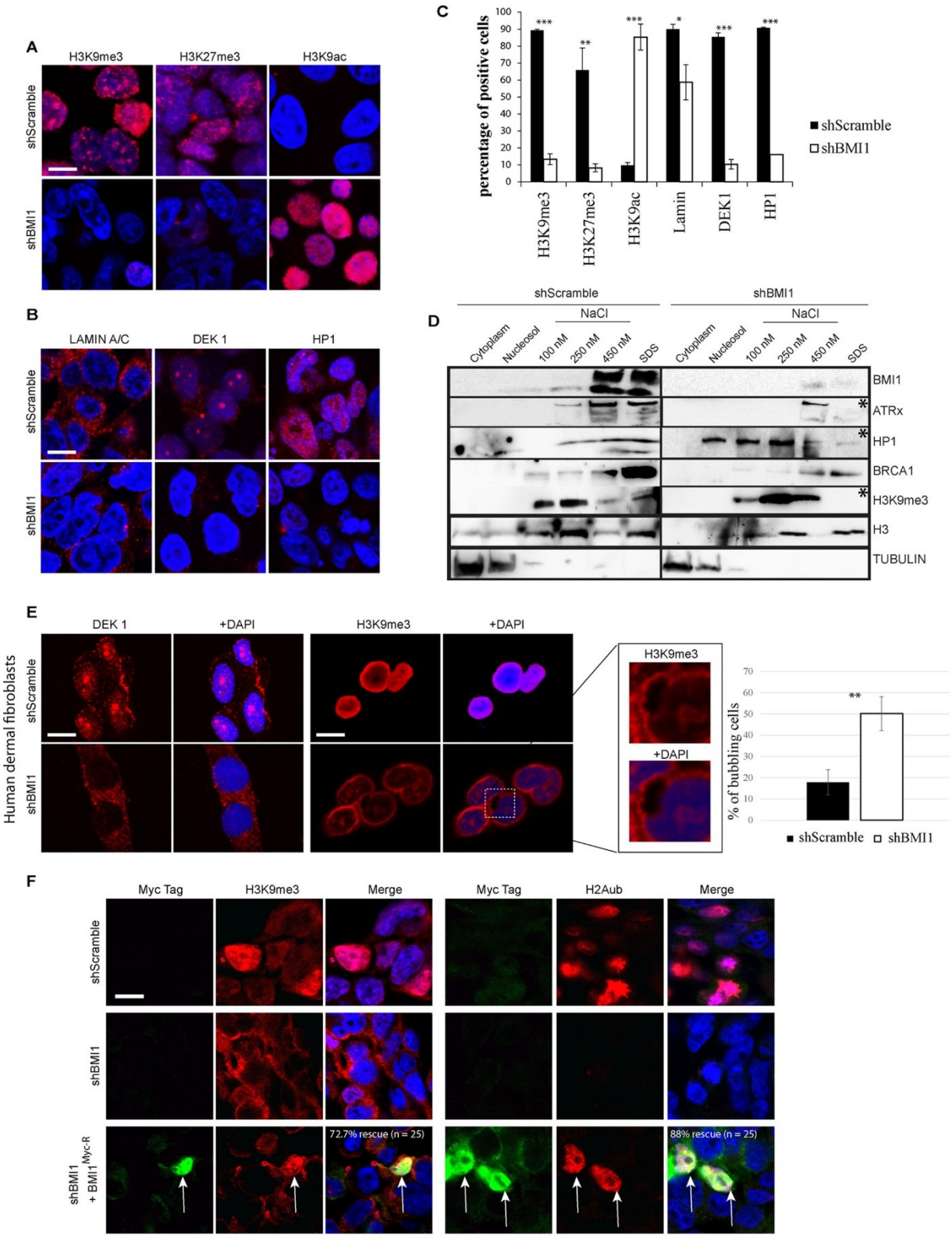


Figure 26. BMI1 knockdown cells present heterochromatin and nuclear envelope alterations

(A-C) Formaldehyde fixed 293FT cells were immunolabeled and counterstained with Dapi. Scale bar, 10µm. Positive cells were counted on 4 different images for a total of 200 cells per condition, and the percentage of positive cells was calculated accordingly. T-test with two tails, where $P \leq 0.05$*, ≤0.01**, ≤0.001***. Note that the apparent localization of LAMIN A/C in the cytosol is the result of Triton X-100 treatment. (D) 293T cells were infected with shScramble or shBMI1 viruses and cell's compartments were fractionated. Note ATRx, HP1 and H3K9me3 reduction (*) in the SDS fractions of shBMI1-treated cells. (E) Human dermal fibroblasts were infected with shScramble or shBMI1 viruses, immunolabeled and counterstained with DAPI. Note reduced DEK1 and H3K9me3 labeling, and H3K9me3 localization at the nuclear periphery, in BMI1 deficient cells. Bubbling of the nuclear envelope was also observed (inset), where $P \leq 0.01$**. Scale bar, 10 µm. (F) Human dermal fibroblasts were infected with shScramble or shBMI1 viruses, and next transfected with a plasmid encoding an RNAi-resistant BMI1 Myc-tagged construct. Note the rescue of H2Aub and H3K9me3 nuclear labeling in Myc-positive cells knockdown for BMI1 (arrows). Scale bar, 10 µm.

2.4.6 BMI1 and BRCA1 display partial functional redundancy in heterochromatin compaction

Since BMI1/RING1A/RING1B and BRCA1/BARD1 display H2A mono-ubiquitination activities and that BRCA1 enrichment and distribution at heterochromatin is not affected upon BMI1 depletion, we tested whether BMI1 and BRCA1 displayed functional redundancy. For this, we first inactivated BRCA1 to test the impact on BMI1 localization. In control cells, both BRCA1 and BMI1 were enriched at repeat-DNA sequences (Fig. 44). In BRCA1 knockdown cells, BRCA1, HP1, H2Aub and H3K9me3 levels were reduced and the transcription of repeat-DNA sequences was increased (Fig. 44C), altogether confirming previous findings (684). Notably however, BMI1 and RING1B were enriched at all tested regions upon BRCA1 knockdown (Fig. 44). Next, stably infected shBMI1 cells were transfected with an shBRCA1 plasmid, generating double knockdown (DKN) cells. While H3K9me3 enrichment was reduced by 55-70% in DKN cells at all tested chromatin regions, enrichment for HP1 and H2Aub was further reduced by 80-90% (Fig. 27A), suggesting additive effects in DKN cells when compared to single BRCA1 or BMI1 knockdown cells. To test whether BMI1 could compensate for BRCA1 deficiency, we over-expressed the $BMI1^{Myc}$ construct. In

control cells, BMI1 over-expression could not displace endogenous BRCA1 localization on the chromatin (Figs. 27B). However, chromatin accumulation of both endogenous and ectopic BMI1 proteins was highly increased at all tested regions following BRCA1 knockdown (Fig. 27B). A similar but less dramatic trend was also observed for RING1B. Most notably, while BMI1 over-expression could increase H2Aub and H3K9me3 deposition as well as HP1 accumulation at all tested regions in control cells, it could also rescue the corresponding heterochromatin anomalies in shBRCA1 cells (Fig. 27B). BMI1 over-expression in shBRCA1 cells also resulted in normalization of *ALU*, *McBox*, *Sata* and *SatIII* expression (Fig. 27C), altogether suggesting functional redundancy in constitutive heterochromatin compaction and silencing between BMI1 and BRCA1.

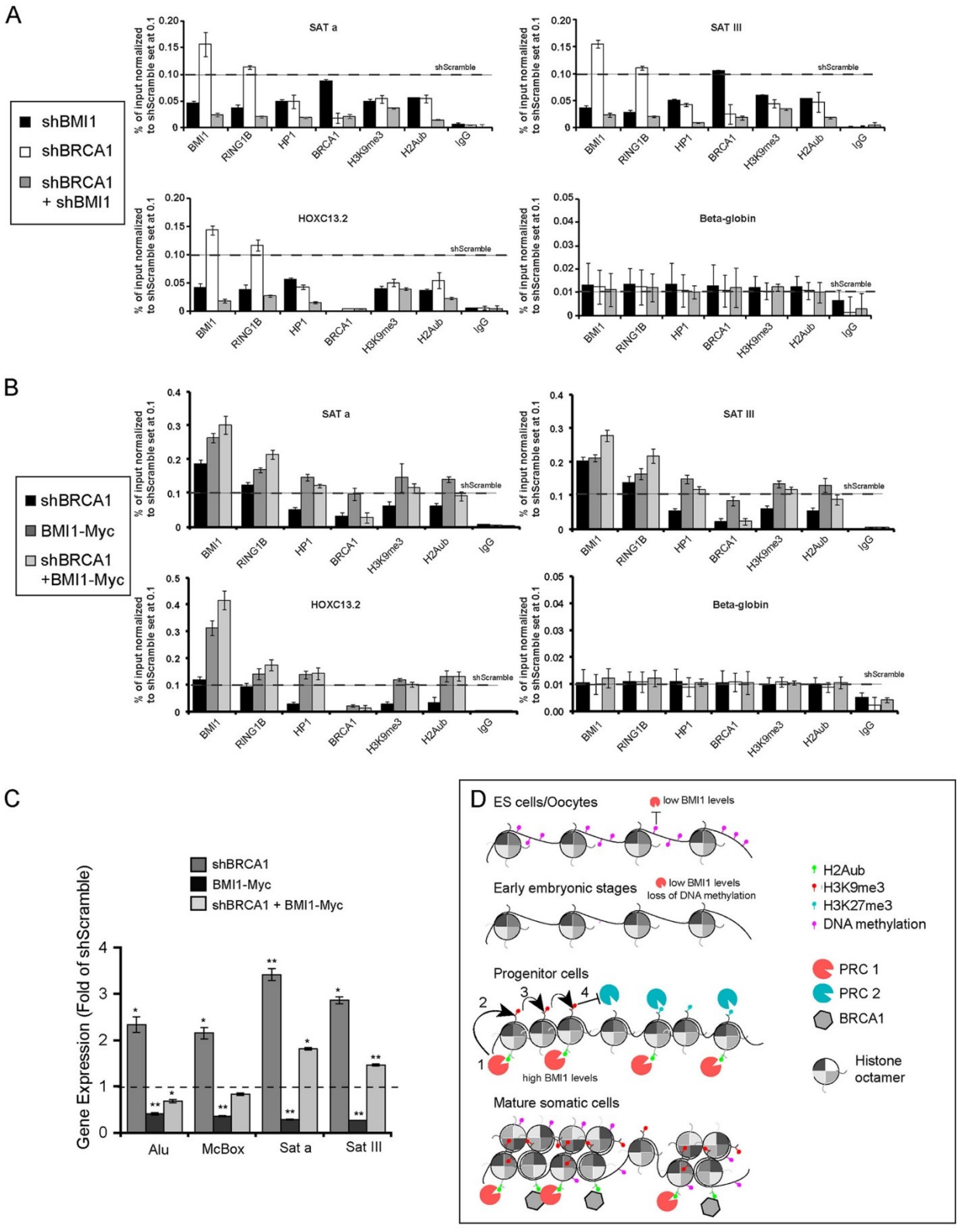


Figure 27. BRCA1 and BMI1 display redundant activities in constitutive heterochromatin formation and silencing

(A) 293T cells were infected with shScramble or shBMI1 viruses. After selection with hygromycin, cells were transfected or not with an shBRCA1-encoding plasmid and analyzed by ChIP. Note the severe reduction for HP1, H3K9me3 and H2Aub at human Satellite repeats in shBMI1/shBRCA1 cells. All data where normalized to shSramble (black horizontal bars). (B) 293T cells stably expressing $BMI1^{Myc}$ or not were transfected with shScramble or shBRCA1 plasmids and analyzed by ChIP (B), and qPCR (C). (B) Endogenous and exogenous BMI1 was enriched in shBRCA1-treated cells at all tested loci. BMI1 over-expression also rescued HP1, H3K9me3 and H2Aub depletion in BRCA1 knockdown cells at human Satellite repeats. All data where normalized to shSramble (black horizontal bars). (C) BMI1 over-expression rescues repeat-DNA sequences expression in BRCA1 knockdown cells, where $P \leq 0.05$*, ≤0.01**. (D) Model depicting the dynamic and developmentally regulated PRC1 (BMI1) occupancy at PCH. In ES cells and oocytes, BMI1 is expressed at low levels and binding to PCH is prevented by DNA methylation. At pre-implantation stages, DNA methylation is erased coincidently with low levels of BMI1. During development, high BMI1 levels in progenitor cells stimulate PCR1-mediated H2Aub (1), which promotes H3K9me3 deposition (2), and propagation (3). In turn, H3K9me3 prevents PRC2-mediated H3K27me3 activity on nucleosomal histones (4). PRC2 occupancy at PCH prior to PRC1 would allow deposition of the H3K27me3 mark (right) and the bivalent histone signature. In mature somatic cells, the PRC1 and BRCA1/BARD complexes are highly enriched at PCH, leading to heterochromatin compaction and silencing. DNA methylation at PCH may occur after spreading of the $H3K9^{me3}$ mark.

2.5 Discussion

We showed here that *Bmi1*-deficient mouse neurons and BMI1 knockdown human cells displayed severe anomalies at the constitutive heterochromatin. These anomalies were accompanied by transcriptional activation of repeat-DNA sequences and correlated with robust accumulation of BMI1 at constitutive heterochromatin. Genome-wide analysis of BMI1 distribution on the chromatin further revealed predominant enrichment at repetitive DNA sequences. BMI1 co-purified with architectural heterochromatin proteins, co-localized with H3K9me3, and was required for HP1, DEK1 and ATRx localization at constitutive heterochromatin. In contrast, BRCA1 localization was BMI1-independent, and both proteins displayed partial functional redundancy for H2Aub deposition, heterochromatin formation and silencing.

The recruitment mechanisms of PcG proteins are complex and not fully understood. In mouse ES cells and oocytes, evidence suggests that recruitment of PcG proteins at PCH is prevented by high level of DNA methylation (374,705). Likewise, observations of Polycomb bodies at PCH in transformed human cells are thought to occur following loss of DNA methylation (374,701,705). Based on this, it was proposed that PcG proteins are excluded from PCH in normal somatic cells (705). Using several methods, we demonstrated BMI1 enrichment at PCH and other repetitive elements in mouse neurons, human neural precursors and immortalized human cells. Our cell fractionation assays further demonstrated that about 50% of the BMI1 pool was bound to the SDS-soluble chromatin fraction, which is enriched for constitutive heterochromatin. Interestingly, we also observed that: 1) BMI1 did not co-purify with EZH2 or H3K27me3; 2) EZH2 and H3K27me3 were not enriched at constitutive heterochromatin; and 3) BMI1 accumulation at constitutive heterochromatin and *HOXC13* was EZH2-independent. Although apparently surprising, these results are in agreement with numerous findings showing that PRC1 recruitment can be PRC2-independent, or that the PRC1 can work upstream of the PRC2 (369,370,373). It is also notable that RING1B knockdown could not mimic the BMI1-deficient phenotype. More specifically, while H2Aub levels were reduced at *HOXC13* in RING1B knockdown cells, this was not accompanied by a corresponding reduction in HP1 and H3K9me3 levels, as observed in BMI1 knockdown cells. Furthermore, there was no apparent effect on H2Aub levels at heterochromatin. This finding leaves us open with many explanations, one being functional compensation by RING1A for H2Aub

deposition, as shown in other context (331,392,703,713). It is also possible that in addition to promote H2Aub at silenced developmental genes and heterochromatin, BMI1 stimulates chromatin compaction and H3K9me3 loading through interactions and activities not shared by RING1A or RING1B (714).

Based on our findings and previously published work (356,700–702), we propose that PRC1 recruitment to PCH is highly dynamic and developmentally regulated (Fig. 27D). The highly variable DNA methylation states between ES cells/oocytes and progenitor cells/somatic cells, in combination with the distinct histone tail modifications and chromatin compactions levels, would explain the re-localization of PRC1 components to PCH in progenitor and somatic cells (715–717). Consistently, BMI1 expression levels are extremely low in human ES cells when compared to human neural progenitors and post-mitotic neurons (V. P., A. B., M. A., A. F. and G. B., unpublished). While PRC1 proteins are excluded from PCH in ES cells and oocytes, they would start to accumulate at PCH during mid-embryonic development coincidently with H3K9me3 deposition and progressive *de novo* DNA methylation (717). This is supported by the ChIP-Seq data showing that BMI1 is moderately enriched at PCH when compared to *HOX* and *p16*^{INK4A} canonical sites in human neural progenitors (Fig. 23). Interestingly, both H3K9me3 and H3K27me3 marks were also present at PCH but did not clearly overlap (Fig. 23), similarly as reported in DNA-methylation deficient mouse ES cells (374). The histone modification pattern of human neural progenitors at PCH is also distinct from that of mouse neurons where H3K27me3 is excluded (Fig. 22). This could be best explained by the robust accumulation of H3K9me3 at PCH in post-mitotic neurons, since H3K9me3 can prevent PRC2, but not PRC1 recruitment (374). Indeed, Bmi1 is highly enriched in mouse neurons at PCH when compared to canonical sites (Fig. 22). Taken together, these observations suggest a dynamic and developmentally regulated model of PcG occupancy at PCH (Fig. 27D). Because constitutive heterochromatin is intrinsically unstable, we further propose that in somatic cells, the main biological function of BMI1 is to stabilize the repetitive genome by promoting chromatin compaction and silencing.

We observed that the *Bmi1*-null neuronal phenotype was associated with increased nuclear diameter and an irregular nuclear envelope. Human cells knockdown for BMI1 also presented anomalies in nuclear envelope architecture (Fig. 26B-E). These anomalies are particularly

interesting considering that loss of heterochromatin foci can result in disruption of the nuclear lamina (670). Perturbations of the nuclear envelope architecture is also a prominent feature of Hutchinson-Gilford Progeria cells carrying mutations in *LAMIN-A* and of normal ageing human cells (548). In most eukaryotes, constitutive heterochromatin perturbations result in genomic instability and premature aging or reduced lifespan (673,679,718–721). It is thus notable that *Bmi1*-deficient mice show reduced lifespan, genomic instability, neurodegeneration and progeria features (286,423,442,697–699). Similar anomalies were also reported for *ATRx*-deficient mice (681,682). Taken together, this raises the possibility that BMI1 requirement for constitutive heterochromatin formation and silencing could underlie the premature ageing/senescence and genomic instability phenotypes observed in *Bmi1*-null mice and cells.

The BRCA1/BARD1 complex is required for heterochromatin formation and silencing through mono-ubiquitination of H2A at PCH, and the genomic instability phenotype of *BRCA1*$^{-/-}$ cells could be rescued by over-expression of a histone H2A protein fused to an ubiquitin moiety in C terminus (684). How BMI1 or BRCA1-mediated H2Aub deposition at repetitive DNA sequences translates into H3K9me3 loading and heterochromatin spreading is unknown. One possibility is that H2Aub induces allosteric changes in the histone H3 lysine tri-methyltransferases SUV39H1/2 to promote their activity, such as proposed for H2Bub and H3K4 methylation (722,723). This would be consistent with previous observations that PRC1 components can interact with SUV39H1 (356). Alternatively, BMI1 may directly or indirectly regulate the transcription of H3 lysine methyltransferases or demethylases, thus operating in trans. We showed here that co-inactivation of BMI1 and BRCA1 induces more severe heterochromatin anomalies than individual BMI1 or BRCA1 deficiencies, and that BMI1 over-expression could rescue the BRCA1-deficient heterochromatin phenotype. These results suggest that BMI1 and BRCA1 are at least partially redundant for H2Aub deposition at constitutive heterochromatin, although the BRCA1/BARD1 complex targets H2A at lysines 127-129 (683). This would indicate that the commonly used anti-H2Aub antibody recognizes both the H2AK119ub and H2AK127-129ub motifs. The observation that BMI1 and BRCA1 proteins accumulation is mutually independent and that BMI1 levels are increased in BRCA1-deficient cells (and reciprocally) also suggests that both protein complexes possibly bind to very close substrates to catalyze H2A mono-ubiquitination. Structural analyses

have indeed revealed a high degree of conservation between the nucleosome-binding loop of BRCA1 and the corresponding domain of RING1B (724). Notably, the reduced neuronal chromocenter number and size phenotype observed in *Bmi1*-null neurons is about identical to that reported for mouse cortical neurons conditionally deficient for *BRCA1*, thus further supporting our findings (684). Taken together, this suggests that although BMI1 and BRCA1 protein complexes target distinct lysine residues on histone H2A, the resulting biological effects on heterochromatin compaction and silencing are highly similar.

In conclusion, we demonstrated for the first time that BMI1 is highly enriched at intergenic repetitive elements and PCH of the mouse and human genomes in normal somatic cells, and required for constitutive heterochromatin formation and silencing. Since BM1 is also present at PCH in cancer cell lines and that several cancer cells were shown to be sensitive to BMI1 inhibition, this raises the possibility that BMI1 may be important to stabilize the transformed heterochromatic genome. BMI1 function at constitutive heterochromatin may be even more critical in *BRCA1*-deficient tumors, thus opening possibilities for the development of synthetic lethal strategies.

3 Chapter 3 – G-quadruplexes originating from evolutionary conserved L1 elements interfere with neuronal gene expression in Alzheimer's disease

The new field of studying G-quadruplex structures is evolving rapidly. The literature showed so far the importance of these structures because they can alter the transcription of genes. For example, c-*MYC*, if they are in their promoter regions, G4 can also complicate the replication and the DDR process because of the bulge created on the DNA by these structures.

Knowing that these structures are thermodynamically more stable than the DNA double helix, most of the studies in the field focus so far on how these structures are resolved and what are the helicases that are implicated in this process. This article focus on how the cell can protect itself against the formation of said structures. In the second part, the article studied what is the effect of these structures if they are not well regulated in a disease context like Alzheimer's disease.

G-quadruplexes originating from evolutionary conserved L1 elements interfere with neuronal gene expression in Alzheimer's disease

Roy Hanna[1], Anthony Flamier[1,2], Andrea Barabino[1], and Gilbert Bernier[1,3]

[1] Stem Cell and Developmental Biology Laboratory, Hôpital Maisonneuve-Rosemont, 5415 Boul. l'Assomption, Montreal, Canada, H1T 2M4

[2] Present address: Whitehead institute of biomedical research, 455 Main Street, Cambridge, MA 02142

[3] Department of Neurosciences, University of Montreal, Montreal, QC, Canada

***** Running title**:** *G-quadruplexes in Alzheimer's disease*

Correspondence: gbernier.hmr@ssss.gouv.qc.ca

Content:
5 Figures
Word count

Abstract: 147

3.1 Summary

DNA sequences containing consecutive guanines organized in 4-interspaced tandem repeats can form stable single-stranded secondary structures, called G-quadruplexes (G4). Herein, we report that the Polycomb group protein BMI1 is enriched at heterochromatin regions containing putative G4 DNA sequences, and that G4 structures accumulate in cells with reduced BMI1 expression and/or relaxed chromatin, including Alzheimer's disease (AD) neurons. In AD neurons, G4 structures preferentially accumulated in lamina-associated domains, and this was rescued by re-establishing chromatin compaction. ChIP-seq analyses revealed that G4 peaks corresponded to evolutionary conserved Long Interspersed Element-1 (L1) sequences predicted to be transcriptionally active. Hence, G4 structures co-localized with RNAPII, and inhibition of transcription reversed the induction of G4 without affecting chromatin's state, thus uncoupling both components. G4 structures were also associated with perturbed neuronal gene expression and splicing in AD. We conclude that chromatin-mediated inhibition of L1 sequences transcription prevents excessive formation of G4 structures in human neurons.

3.2 Introduction

Nucleosomal histone proteins are regulated by post-translational modifications that can impact transcription, replication, and repair. BMI1 is a component of the Polycomb Repressive Complex 1 (PRC1). The PRC1 promotes chromatin compaction and gene silencing in part through its E3-mono-ubiquitin ligase activity mediated by RING1a/b on histone H2A at lysine 119 (H2Aub) (392,687,725). The PRC1 is recruited to facultative heterochromatin to maintain repression at developmental and senescence-associated-genes (726–730). BMI1 is also enriched at constitutive heterochromatin, where it co-purifies with ATRX, HP1, DEK1, and Lamins (731). BMI1 inactivation in human dermal fibroblasts (HDFs) results in loss of heterochromatin and transcriptional de-repression of repetitive DNA sequences (731). More recently, reduced neuronal expression of BMI1 was associated with late-onset sporadic Alzheimer's disease (AD) (448). Acute *BMI1* inactivation in cultured human neurons can also recapitulate AD-associated hallmarks, including the accumulation of beta-amyloid and hyper-phosphorylated Tau (448). Aged mice hemizygous for *Bmi1* (*Bmi1*$^{+/-}$) also develop, along with some progeroid features, AD-like behavioral and neuropathological phenotypes (447). Furthermore, loss of heterochromatin and genomic instability at repetitive DNA sequences were described as new molecular characteristics present in cortical neurons from both *Bmi1*$^{+/-}$ mice and AD cases (447,660). Loss of heterochromatin and transcriptional activation of specific classes of endogenous retroelements occur in neurodegenerative tauopathies and in animal models of Tau over-expression (732,733). Notably, advanced aging is the most significant risk factor to develop AD (734,735), and many anomalies present in AD patient's neurons *in situ*, such as relaxed heterochromatin and nuclear envelope defects, are also considered as hallmarks of aging (734,736–738).

Interestingly, most genetically inherited progeroid syndromes, such as Werner, Bloom, and Xeroderma pigmentosum, present genomic instability phenotypes. These progeroid disease genes encode DNA damage and/or repair proteins. More specifically, Werner (*WRN*), Xeroderma pigmentosum (*XPB*, *XPD*) and Bloom (*BLM*) gene products encode DNA helicases that can resolve G-quadruplex (G4) DNA's secondary structures (also called structured DNA or G-quadruplexes) stabilized by Hoogsteen hydrogen bonds between guanines (G) (128,135,136). Notably, the DNA-dependent ATPase and helicase ATRX is enriched at repetitive DNA sequences predicted to form

G4 structures, and ATRX can physically bind structured DNA *in vitro* (126). Persistent G4 structures have been proposed to represent a threat to genomic stability and gene function by interfering with fork elongation and DNA repair during replication and transcription (161,175,176,739). In normal physiological conditions, however, G4 structures may be important for the control of gene expression, maintenance of telomeres, and establishment of replication origins (740).

We report here that putative G4 DNA sequences are significantly enriched in BMI1 chromatin immunoprecipitation and sequencing (ChIP-seq) data sets, and that *BMI1*-deficiency in healthy somatic cells resulted in the induction of G4 structures following the relaxation of the heterochromatin. Notably, G4 structures were also found to accumulate in AD neurons *in situ* and *in vitro*, as well as in healthy neurons treated with histone deacetylase inhibitors (HDACi). ChIP-seq analyses of post-mitotic human neurons using the 1H6 antibody further revealed that about 95% of peaks corresponded to “active” evolutionary conserved Long Interspersed Element-1 (L1) sequences. Consistently, inhibition of RNA Polymerase II (RNAPII) could reverse the induction of G4 structures without altering the chromatin compaction state. Importantly, G4 structures present at specific loci were found to be associated with perturbed neuronal gene expression and alternative splicing in AD neurons. These results suggest that chromatin-mediated transcriptional repression of L1 sequences prevents excessive formation of G4 structures in human neurons, which otherwise can interfere with normal gene expression.

3.3 Results

Using public BMI1 ChIP-seq data sets, we annotated all BMI1-enriched chromatin regions and tested their propensity to form putative G4 DNA structures using the Quadparser algorithm (100). Using four different sets of Quadparser parameters, we found that BMI1 was significantly enriched at chromatin regions predicted to form G4 structures independently of the Quadparser stringency (Fig. 28a). We further investigated whether BMI1 association with G4 motifs was comparable to that of proteins known to physically interact with structured DNA such as the ATRX, XPB, and XPD helicases (126,136). We annotated ATRX, XPB and XPD ChIP-seq peaks alongside BMI1 peaks for the presence of putative G4. While ~45% of ATRX, XPB, and XPD peaks were associated with putative G4, only ~6% of BMI1 peaks showed the same association (Fig. 28b), suggesting that BMI1 does not bind G4 structures but is rather enriched at chromatin regions with propensity to generate G4. We also analyzed the ChIP-seq peaks obtained with the BG4 antibody, which recognizes G4 structures. As reported, we found that ~60% of BG4 peaks contained a putative G4 DNA sequence (Fig. 28b) (741). Next, we compared the putative G4 DNA sequences contained in the BG4, XPB, and the BMI1 peaks. Surprisingly, while ~30% of the putative G4 DNA sequences linked by BG4 were enriched for XPB, only 0.05% were enriched for BMI1 (Fig. 28c). Similarly, only 0.15% of the putative G4 sequences bound by XPB were enriched by BMI1 (Fig. 28c). Since BG4 peaks were reported to be predominant at the promoter of actively transcribed genes, these results suggested that BMI1 peaks containing putative G4 DNA sequences are rarely present in actively transcribed regions.

To test the possibility that the BMI1 function is required to prevent the formation of G4 structures, we used the 1H6 and BG4 antibodies, which recognize G4 structures (108,741). In early passage normal HDFs, we noticed that the baseline level of 1H6 and BG4 immuno-reactivity was relatively low (Fig. 28d, Fig 29b). However, we observed a robust nuclear and modest cytoplasmic immunoreactivity for 1H6 and BG4 in *BMI1* knockdown HDFs (sh*BMI1*) (Fig. 28d, Fig 29b), suggesting accumulation of both G4 DNA and G4 RNA structures (107,111). To test for the specificity of the signal obtained, we fixed naive cells with paraformaldehyde and then exposed them to HCl, an agent that denatures DNA's secondary structure. HCL is commonly used in transmission electron microscopy and immunohistochemistry on paraffin sections. We found that

HCL treatment resulted in non-specific immuno-labeling when compared to cells retaining a native chromatin state (Fig. 28d, Fig 29b). This revealed that G4-specific antibodies should only be used on cells with native chromatin. Pyridostatin is an agent that stabilizes spontaneously forming G4 structures, and we found that 1H6 immuno-labeling was significantly increased after exposing HDFs to pyridostatin (Fig. 28e) (142). To further validate our observations, we performed co-localization studies with an antibody against the Werner (WRN) helicase. WRN is predicted to bind and unwind G4 structures (130,742,743). In control HDFs, WRN and 1H6 levels were low and did not co-localize (Fig. 28f and Fig. 29c). They were, however, significantly induced after *BMI1* knockdown or after exposition to pyridostatin (Fig. 29c). In both cases, 1H6 and WRN presented a very high coefficient of co-localization (Pearson correlation: 0.79 for sh*BMI1*; 0.70 for shScr + pyridostatin) (Fig. 28f). Relatively strong co-localization between 1H6 and XPB or 53BP1 (but not XPD) was also observed (Fig. 28f and Fig. 29c). Importantly, G4 structures were not detected with 1H6 after treatment of HDFs with 10Gy of gamma radiations, which induce DNA double-strand breaks (Fig. 28g) (744). Taken together, these experiments suggested that nuclear foci visualized using the 1H6 antibody represent genuine G4 structures.

Given that BMI1 is enriched at heterochromatin, and that BMI1 deficiency results in loss of heterochromatin compaction (731), we reasoned that BMI1 activity may counteract the induction of G4 structures through chromatin compaction. The H3K9me3 histone mark is enriched at constitutive heterochromatin (closed) regions, while the H3K9ac histone mark is enriched at euchromatin (open) regions. Accordingly, loss of H3K9me3 nuclear foci was observed in HDFs 16h after transfection with a *BMI1*-targeting shRNA vector (Figs. 28h and 30a). Notably, we found that nearly all 1H6 foci induced following *BMI1* knockdown did not co-localize with H3K9me3 (Fig. 28h, Pearson coefficient correlation of -0.135). In contrast, H3K9ac signal intensity was increased upon *BMI1*-knockdown (Figs. 28h and 30a), and a significant positive correlation (Pearson coefficient correlation of 0.25) was observed between 1H6 and H3K9ac labeling (Fig. 28h) (745). This suggested that the induction of G4 structures may be associated with chromatin relaxation. To test our hypothesis, we used HDACi, which leads to chromatin relaxation by preventing the deacetylation of histones (745,746). We found that HDFs treated with Sodium Butyrate (SB) or Trichostatin A (TSA) displayed rapid induction of G4 structures within 2h, which was markedly

preceded by robust elevation of H3K9ac levels (Figs. 28i and 30 b-d). Pearson correlation analyses at 2h revealed a near-perfect correlation between H3K9ac and 1H6 labeling, suggesting that most G4 structures were induced following histone acetylation (Figs. 28i and 30 b-d).

To test if *Bmi1* knockout was also associated with the formation of G4 structures, we analyzed *Bmi1*$^{-/-}$ mice. The Bmi1 protein is expressed in post-mitotic retinal neurons (Fig 31a), including cone photoreceptors. Cones from *Bmi1*$^{-/-}$ mice present reduced heterochromatin compaction at post-natal (P) day 25 (286,443). In wild type (WT) mice, we found that the baseline level of 1H6 was low in all retinal neurons (Figs. 28j and 31a). In contrast, nuclear 1H6 immunoreactivity was high in S-opsin positive cones of *Bmi1*$^{-/-}$ mice (Figs. 28j-white arrows and 31a). Notably, 1H6 induction in *Bmi1*$^{-/-}$ cones also correlated with an increased H3K9ac level (Fig. 31b-white arrows). Interestingly, 1H6 was also induced in cones from *Bmi1*$^{-/-}$ mice at P1 (Fig. 31c), thus before the onset of retinal degeneration (443). These results suggested that chromatin compaction prevents excessive formation of G4 structures in both human and mouse cells.

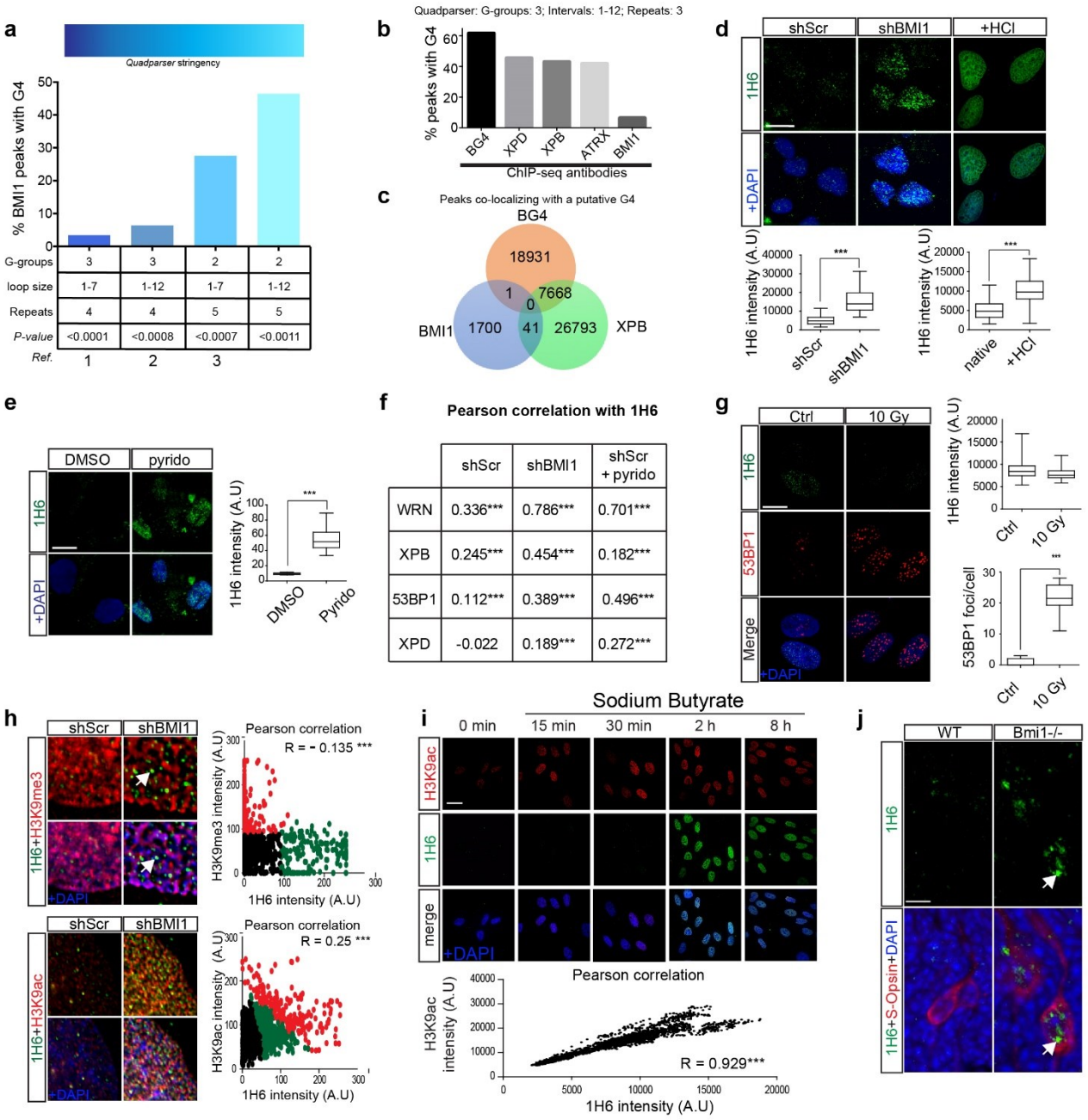


Figure 28. Chromatin compaction prevents excessive formation of G4 structures

a. Showed is the proportion of BMI1 ChIP-seq peaks containing a putative G-quadruplex motif according to the Quadparser algorithm and using four independent sets of parameters. Top: color gradient indicating the stringency of the Quadparser parameters. G4, G-quadruplex; 1 Gray et al. 2014 parameters (747); 2 Gray et al. 2014 and Zizza et al. 2016 (178); 3 Law et al. 2010 (126); P-

value: probability value based on generating six sets of 3,542 randomly positioned probes and annotated for G4 motifs.

b. The proportion of ChIP-seq peaks for XPB, XPD, ATRX, and BMI1 containing a putative G-quadruplex motif according to the Quadparser algorithm. Gray et al. parameters were used to annotate all ChIP-seq datasets. G4, G-quadruplex.

c. Venn diagram for BMI1, XPB, and BG4 (G4-seq) chip-seq peaks co-localizing with a putative G-quadruplex according to Gray et al. Quadparser parameters.

d. Formaldehyde fixed HCA2 cells infected with shScramble or shBMI1 viruses, or denaturated using 3M of HCl, were immunolabeled and counterstained with DAPI and 1H6 antibody. These antibodies were used to detect G-quadruplexes structures. The graphs show the quantification of 1H6 signal intensity in each cell with the relevant t-test. P ≤ 0.05*, ≤0.01**, ≤0.001***. Scale Bar: 10 µm

e. Formaldehyde fixed HCA2, treated for 16h with vehicle (DMSO) or 5 µM of pyridostatin (pyrido), were immunolabeled and counterstained with DAPI. The 1H6 antibody was used to detect G4 structures. Scale Bar: 10 µm. The graphs show the quantification of 1H6 signal intensity in each cell from Fig. 1E with the relevant t-test. P ≤ 0.05*, ≤0.01**, ≤0.001***.

f. Pearson correlation between 1H6 and various helicases co-localization was calculated from the immunofluorescence pictures (see Figure 29C).

g. 10 Gy irradiated HCA2 cells were fixed then immunolabeled and counterstained with DAPI. The graphs show the quantification of 1H6 and 53BP1 signal intensity in each cell. An average of 150 cells was analyzed per condition with the relevant t-test. P ≤ 0.05*, ≤0.01**, ≤0.001***. Scale Bar: 10 µm

h. Formaldehyde fixed HCA2 cells infected with shScramble or shBMI1 viruses were co-immunolabeled with 1H6 and H3K9me3 or H3K9ac antibodies and counterstained with DAPI. Pearson correlation between 1H6 and various histone modifications co-localization was calculated from the immunofluorescence pictures. The scatter plot was then divided by a K-means clustering using 3 groups, and the Pearson coefficient is indicated on the graph. Scale Bar: 10 µm.

i. Immunofluorescence pictures of HCA2 cells treated with 10 µmol/mL SB for the indicated time. Pearson correlation study of the co-expression between 1H6 and H3K9ac at the 2h time point. Plotted in a scatter graph with the Pearson coefficient indicated on the graph. An average of 150 cells was analyzed per condition. See more quantification in Figure 30 D Scale Bar: 40 µm

P ≤ 0.05*, ≤0.01**, ≤0.001***. All values are means ± SEM.

j. IF analyses on WT and *Bmi1*$^{-/-}$ mouse retinal sections at P10 using the 1H6 and anti-S-Opsin antibodies. S-cone photoreceptors with the induction of G4 are showed (white arrows). Scale bars: 12µm

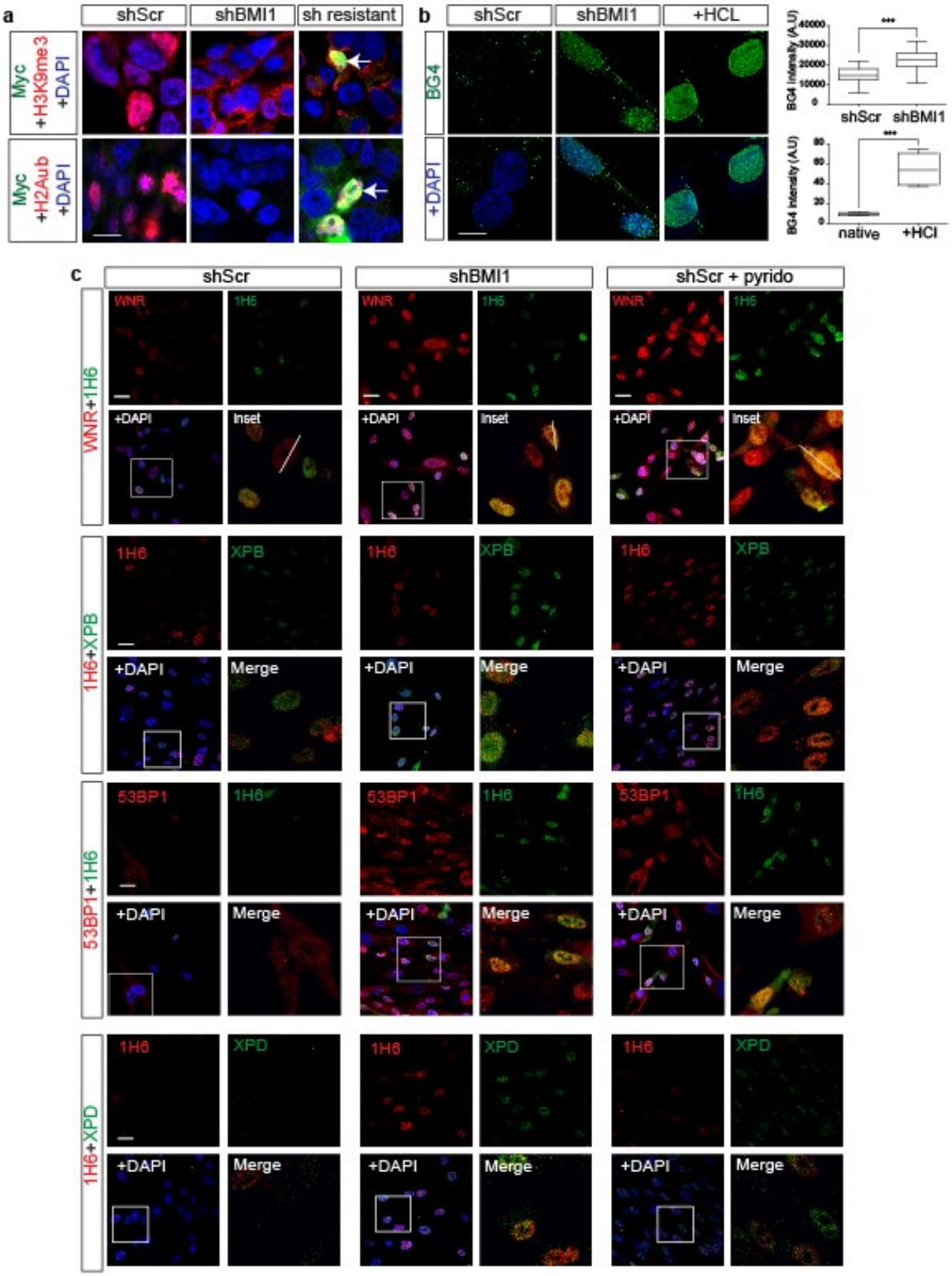


Figure 29. Validating the 1H6 antibody as a suitable antibody to detect G4 structures

a. HCA2 cells were infected with the shScramble or shBMI1 viruses together with a BMI1myc DNA construct that was shRNA-resistant. 72 hours later, paraformaldehyde-fixed cells were immunolabeled and counterstained with DAPI. Cells positive for the BMI1myc construct are indicated (white arrow). Note the rescue of H2Aub and H3K9me3 in shBMI1 cells, also expressing the shRNA-resistant BMI1myc construct. Scale bar: 15µm.

b. Formaldehyde fixed HCA2 cells infected with shScramble or shBMI1 viruses, or denaturated using 3M of HCl, were immunolabeled and counterstained with DAPI and BG4 antibody. These antibodies were used to detect G-quadruplexes structures. The graphs show the quantification of BG4 signal intensity in each cell with the relevant t-test. $P \leq 0.05$*, ≤0.01**, ≤0.001***. Scale Bar: 10 µm

c. Formaldehyde fixed HCA2 cells were co-immunolabeled with 1H6 and antibodies against various helicases, counterstained with DAPI. Crop with higher magnification of the area is indicated by the square. Scale Bar: 10 µm

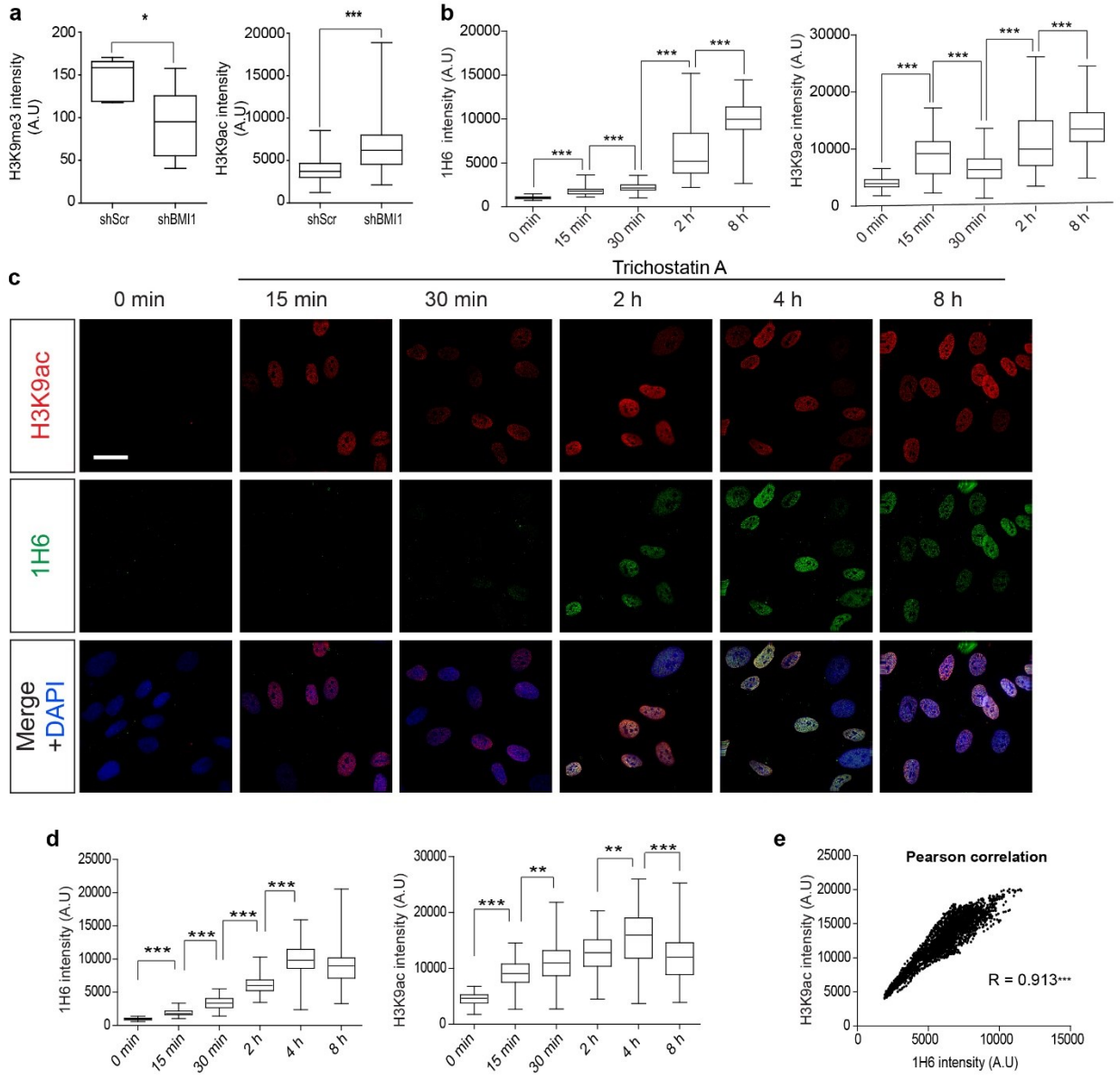


Figure 30. Chromatin relaxation induces the formation of G4 structures

a. Quantification of the H3K9me3 and the H3K9ac signal from Figure 28I. BMI1 knockdown resulted in decreased H3K9me3 level and increased H3K9ac level. Statistical differences were analyzed using unpaired T-test with two tails.

b. Quantification of the H3K9ac and 1H6 nuclear signal in the cells that were treated with sodium butyrate in Figure 28J. Statistical differences were analyzed using unpaired T-test.

c. Immunofluorescence pictures of HCA2 cells treated with 5ng/mL of trichostatin A for the time indicated. Scale bar: 35 µm.

d. Mean signal intensity of 1H6 and H3K9ac in each nucleus were measured and plotted in a box and whisker graph. Statistical differences were analyzed using unpaired T-test with two tails.

e. Pearson correlation study of the coexpression between 1H6 and H3K9ac at the 2h time point plotted in a scatter graph with the Pearson coefficient indicated on the graph. P ≤ 0.05*, ≤0.01**, ≤0.001***.

µm. P ≤ 0.05*, ≤0.01**, ≤0.001***.

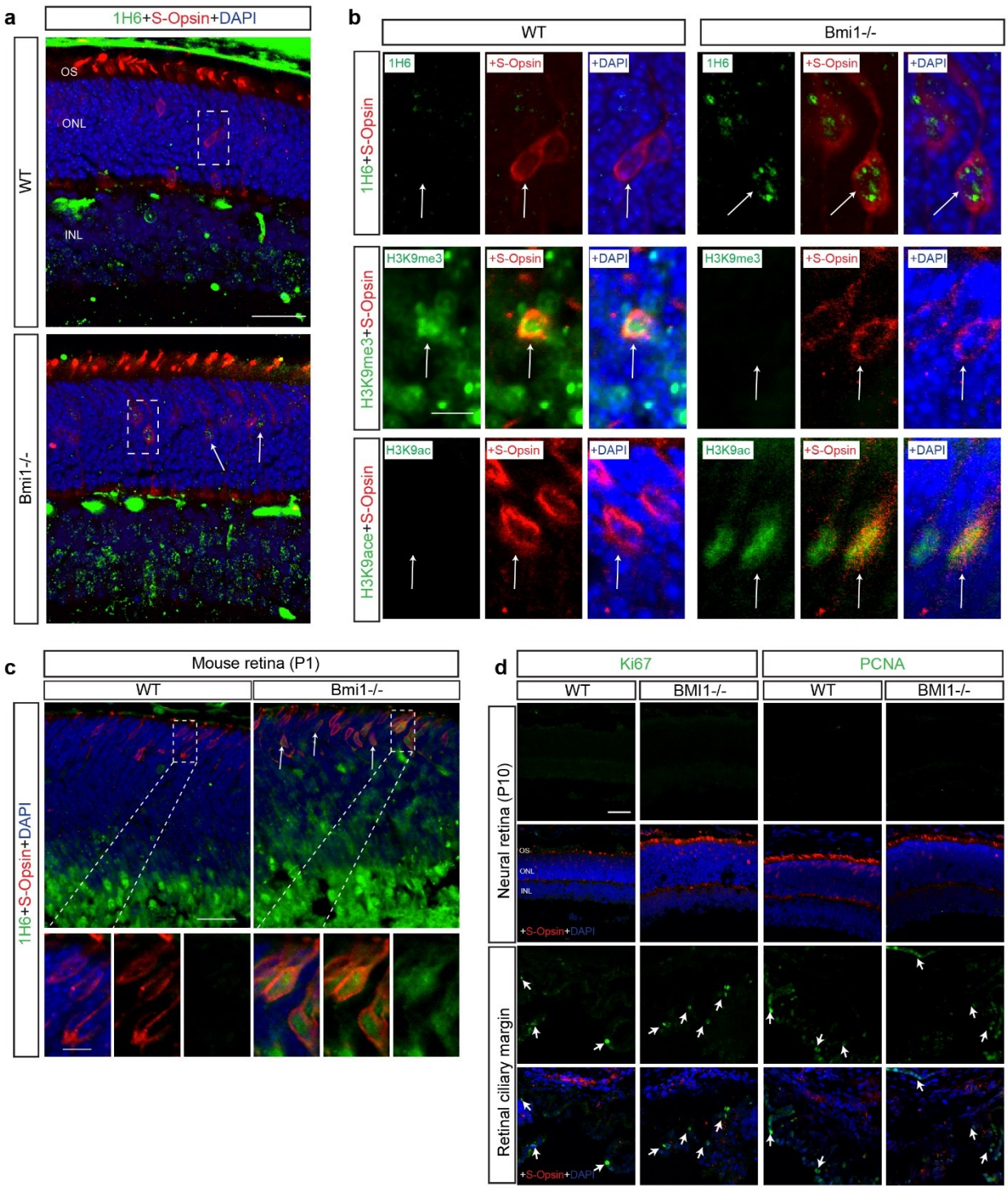


Figure 31. Loss of Bmi1 leads to G4 structures in mouse photoreceptors.

a. IF analyses on WT and *Bmi1*$^{-/-}$ mouse retinal sections at P10 using the 1H6 and anti-S-Opsin antibodies. Crop with higher magnification of the area indicated by the respective dashed rectangles are presented in Figure 1.E. Scale bars: 40µm

b. IF analyses on WT and *Bmi1*$^{-/-}$ mouse retinal sections at P10 using 1H6, anti-H3K9ac, anti-H3K9me3, and anti-S-Opsin antibodies. Photoreceptors were labeled with H3K9ac or H3K9me3 antibodies (white arrows). Scale bars: 12µm

c. IF analyses of WT and *Bmi1*$^{-/-}$ mouse retinas at P1 using 1H6 (G4) and anti-S-Opsin antibodies. Crop with higher magnification of the area indicated by the respective dashed rectangles (at the bottom). S-cone photoreceptors with the induction of G4 structures are shown (white arrows). Scale bars: 40µm (top); 12µm (at the bottom)

d. IF analyses of WT and *Bmi1-/-* mouse retinas at P10 using the cell cycle markers anti-Ki67 and anti-PCNA and anti-S-Opsin (s-cone photoreceptors marker) antibodies. Positive control staining: positive cells for Ki67 and PCNA staining in the retinal ciliary margin (white arrows). Scale bars: 40µm

3.3.1 Re-establishing chromatin compaction reverses accumulation of G4 structures in AD neurons

Considering that *BMI1* expression is reduced in AD brains and neurons (448), we tested whether the above findings were relevant to AD. We first evaluated the presence of G4 structures in *BMI1* knockdown human neurons, an experimental cellular model of AD (448). We found that *BMI1* knockdown neurons showed reduce H3K9me3 levels, which correlated with the robust accumulation of G4 structures (Fig. 32a). Using frozen brain sections from the hippocampus of elderly controls and AD cases, we found that G4 structures also accumulated in AD neurons *in situ* (Fig. 32b). Likewise, iPSC-derived neurons from two distinct AD cases showed increased H3K9ac levels that correlated with the accumulation of G4 structures (Figs. 32c, 33a, b). DNA damage accumulates in AD neurons *in situ* (447), and 53BP1 nuclear foci label DNA damage (748). By performing a time-course study, we observed that AD neurons showed an accumulation of G4

structures at day in *vitro* (DIV) 14, thus before the appearance of 53BP1 nuclear foci at DIV30 (Fig. 32d). 1H6 also did not co-localized with 53BP1 (not shown), suggesting that G4 structures present in AD neurons are not associated with DNA damage.

We performed high-resolution confocal microscopy and 3D reconstruction analyses on AD neurons to reveal the subcellular localization of G4 structures. G4 structures were found to be abundant around the nucleolus and at interspaced puncta located at the nuclear periphery, suggesting accumulation at the nucleolar heterochromatin and at lamina-associated domains (LADs), respectively (Fig 32e) (749,750). Using the neuronal marker βIII-tubulin, we also confirmed that G4 structures observed in AD cultures were predominant in neurons (Fig. 33a). We further investigated if the modulation of chromatin structure could improve the observed phenotype. Tau over-expression in neurons is sufficient to induce heterochromatin relaxation, and GSK3b is the primary kinase that phosphorylates Tau in AD (660,751,752). Likewise, p53 accumulates in AD neurons, and p53 can initiate heterochromatin relaxation by inhibiting the expression of the histone Lys9 tri-methyltransferase *SUV39H1* (448,753). We thus treated control and AD neurons for 24h with CHIR99021 (a GSK3b inhibitor) or pifithrin-alpha (a p53 inhibitor) (Fig. 32f). We found that both treatments significantly reduced H3K9 hyper-acetylation and the accumulation of G4 structures in neurons from two unrelated AD cases (Figs. 32f and 33b-d), revealing that re-establishment of chromatin compaction is sufficient to rescue the G4 phenotype.

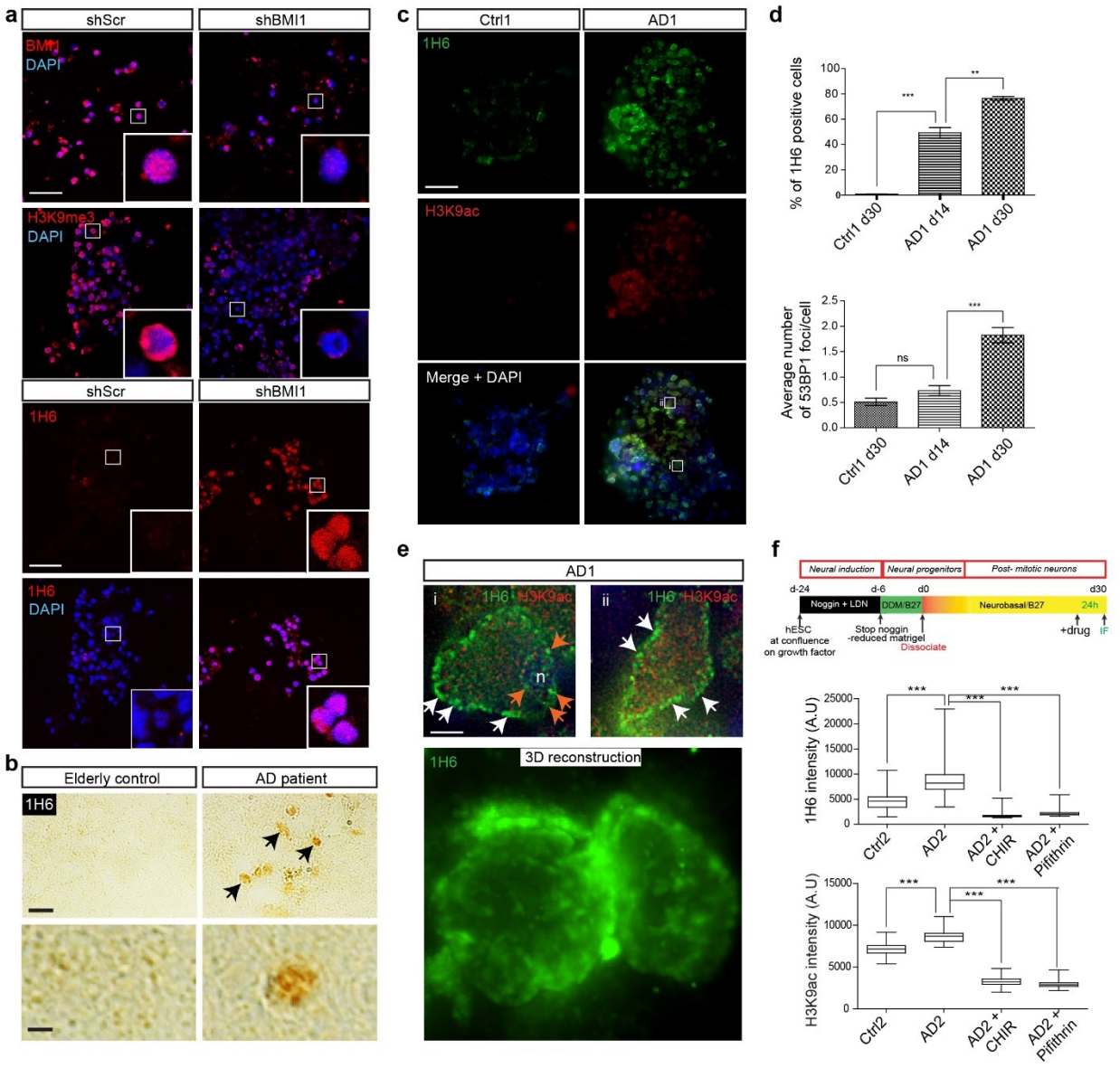


Figure 32. Re-establishing chromatin compaction reverses the accumulation of G4 structures in AD neurons

a. IF analysis showing that BMI1 knockdown in human neurons results in loss of heterochromatin (H3K9me3) and induction of G4 DNA structures (1H6). Scale bar: 35 µm.

b. IHC analysis on frozen human brain sections (frontal cortex) showing 1H6 immunoreactivity in AD patient's neurons.

c. IF analysis showing G4 DNA structure (1H6) induction and histone H3 acetylation (H3K9ac) in AD neurons. Scale bar: 35 µm.

d. Quantification of IF results showing G4 DNA induction in AD neurons that occurred before 53BP1 accumulation. Statistical differences were analyzed using an unpaired T-test with two tails, and the result is shown on the graph. P ≤ 0.05*, ≤0.01**, ≤0.001***.

e. High-magnification IF analysis showing G4 DNA structures at the nuclear membrane (white arrows) and nucleolar (n) periphery (orange arrows) in AD neurons. This phenotype can also be visualized using 3D reconstruction. Scale bar: 10 µm.

f. Schematic of the method used to produce iPSC-derived neurons (top image). Control (Ctrl) and AD iPSC-derived cortical neurons were treated for 24h with an inhibitor of GSK3β (CHIR99021) or an inhibitor for p53 (pifithrin). They were then labeled with the 1H6 and H3K9ac antibodies and counterstained with DAPI prior to immunofluorescence analysis. Mean 1H6 or H3K9ac fluorescence intensity/cell was quantified and plotted in a whisker and box plot. Statistical differences were analyzed using an unpaired T-test with two tails. P ≤0.001***. All values are means ± SEM.

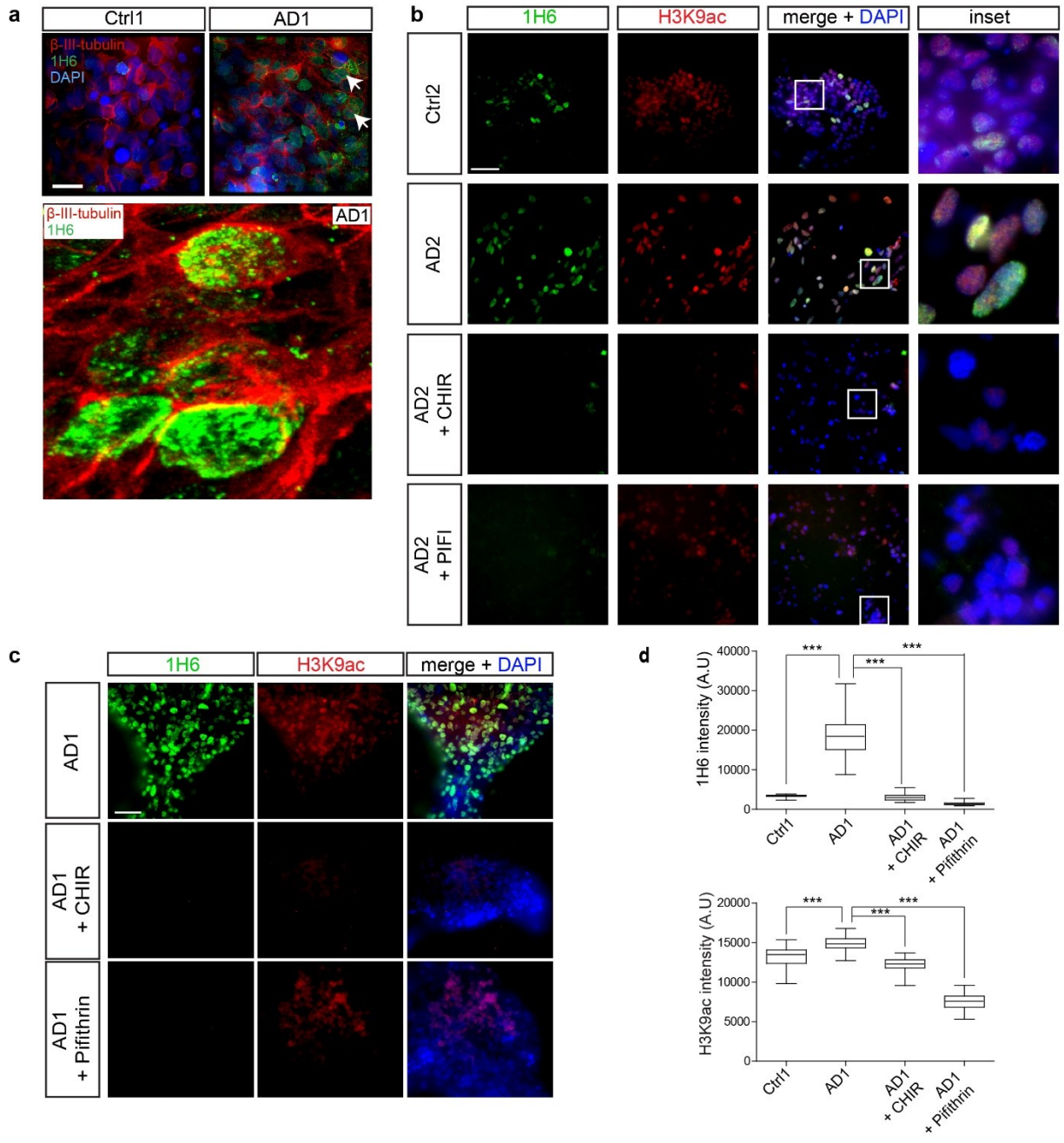


Figure 33. The formation of G4 structures can be reversed by re-establishing chromatin compaction

a. IF analysis showing that 1H6-positive cells in AD cultures are βIII-tubulin-positive neurons. Scale bar: 35 µm.

b. IF analyses of iPSC-derived control (Ctrl) and AD2 cortical neurons. Mature neurons were treated for 24h with an inhibitor of GSK3β (CHIR99021) or an inhibitor for p53 (pifithrin), then labeled with antibodies against 1H6 and H3K9ac. Scale bar: 34 µm.

c. IF analyses of iPSC-derived control (Ctrl) and AD1 cortical neurons. Mature neurons were treated for 24h with an inhibitor of GSK3β (CHIR99021) or an inhibitor for p53 (pifithrin), then labeled with antibodies against 1H6 and H3K9ac. Scale bar: 34 µm.

d. Mean 1H6 or H3K9ac fluorescence intensity/cell was quantified and plotted in a whisker and box plot. Statistical differences were analyzed using an unpaired T-test with two tails. P ≤0.001***. All values are means ± SEM.

3.3.2 1H6 peaks present high similarities with canonical and non-canonical G4 sequences

In order to localize G4 structures on the human genome, we performed ChIP-seq using the 1H6 antibody on control and AD neurons at DIV30 (448). We identified 1389 peaks in AD neurons and 1165 peaks in control neurons. Enrichment of 1H6 within the peaks was significantly higher in AD neurons (Figs. 34a and 35a). Peaks in AD neurons were also significantly larger than in controls, covering 0.6% of the human genome, compared to 0.2% (Fig. 34b). While most of the peaks in control neurons were represented in AD samples (Fig. 35a), cluster 1 from the AD peaks showed no enrichment in control neurons, revealing a large subgroup of AD-specific peaks (Fig. 34a). Venn diagram distribution revealed that out of the 1389 AD peaks, 737 peaks were AD-specific, and 558 were shared with control neurons. Out of the 1165 control peaks, 504 peaks were control-specific, and 609 were shared with AD neurons (Fig. 34c). The discrepancy in the number of peaks shared by both groups was explained by the occasional presence of two control peaks within a very large and unique AD peak.

We observed by immuno-fluorescence that most G4 structures induced in AD neurons localized at LADs. Bioinformatic analysis of ChIP-seq data further revealed that ~40% of 1H6 peaks colocalized with constitutive LAD domains (cLADs) and ~25% with facultative LAD domains (fLADs)

(Fig. 34e) (754). AD peaks were also found to colocalized more with cLADs than control peaks. Hence, ~65% of AD peaks indeed colocalized with cLADs or fLADs. The observed percentage could not be explained by the coverage of these domains on the genome since cLADs, and fLADs represent 49.7% of the human genome (Fig. 34f).

Next, we compared all 1H6 peaks with the canonical G4 motif. The most stringent predicted canonical G4 sequence contains 4 interspaced repeats with at least 3 guanines/repeat, the guanine repeats being separated by a loop of 1-12 nucleotides (93). This revealed that 61% of 1H6 peaks in AD neurons and 53% of 1H6 peaks in control neurons colocalized with a predicted canonical G4 sequence (Fig. 35b). Importantly, these values presented a Z-score nine times higher than a random distribution of the peaks on the human genome (Fig. 35c). Annotation of all 1H6 peaks revealed that ~38% were located within gene bodies, ~20% at enhancers, and ~5% at CpG islands (Fig. 34d). We also analyzed all peaks using an unbiased motif discovery algorithm. We found that the most statistically significant motifs discovered by MEME Suite contained interspaced G repeats predicted to form non-canonical G4 structures (Figs. 34g, h, and 35e) (755,756). When probed with a QGRS finder (Quadruplex forming G-Rich Sequences) (757), these motifs presented a G-score of 34 (758). When analyzing with MEME the peaks that did not colocalize with a canonical G4 sequence (~45% of all peaks), we found that the most statistically significant motifs also contained interspaced G repeats predicted to form non-canonical G4 structures (Fig. 35f). These results indicate that most epitopes recognized by the 1H6 antibody in the ChIP-seq experiment corresponded to structured DNA.

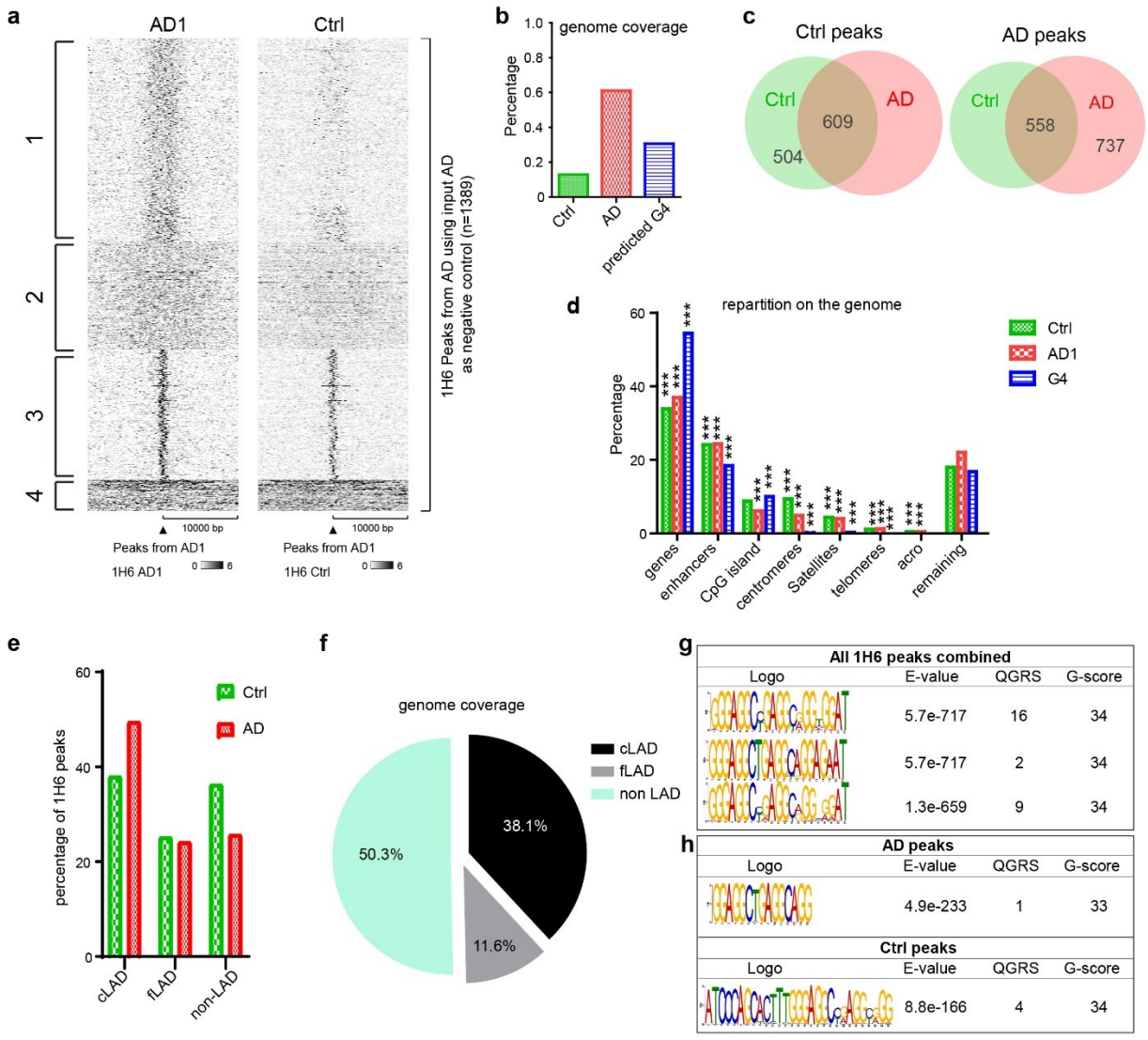


Figure 34. 1H6 peaks present high similarities with canonical and non-canonical G4 sequences

a. Heatmap of 1H6 ChIP-seq enrichment from AD neurons or control neurons centered on AD peaks with a +/- 10 Kb. K-means clustering highlighted on the right was done using the AD enrichment.

b. Histogram showing the percentage coverage of 1H6 peaks and predicted G4 sequences on the entire human genome.

c. Venn diagram shows the number of specific AD peaks and those that are shared with a Ctrl peak and a diagram that shows the Ctrl peaks that are specific and those that are shared with an AD peak.

d. Repartition of 1H6 peaks and predicted G4 sequences on the genome. A permutation test was performed for each combination in order to calculate the significance. A total of 1000 permutated sets of probes were randomly generated and annotated for the different parts of the genome.

e. graph showing the percentage of colocalization between 1H6 peaks and LAD domains or non LAD domains.

f. A pie chart showing the repartition of the genome between LAD domains and non LAD domains.

g. MEME analysis on all the ChIP-seq combined on the neurons showing the top three motifs E-Value associated with them, the number of QGRS (Quadruplex forming G-Rich Sequences) and the highest G-score associated with these sequences.

h. MEME analysis on the 1H6 peaks of Ctrl or AD neurons showing motifs that can form a G4 structure with the E-Value associated with them, the number of QGRS, and the highest G-score associated to these sequences.

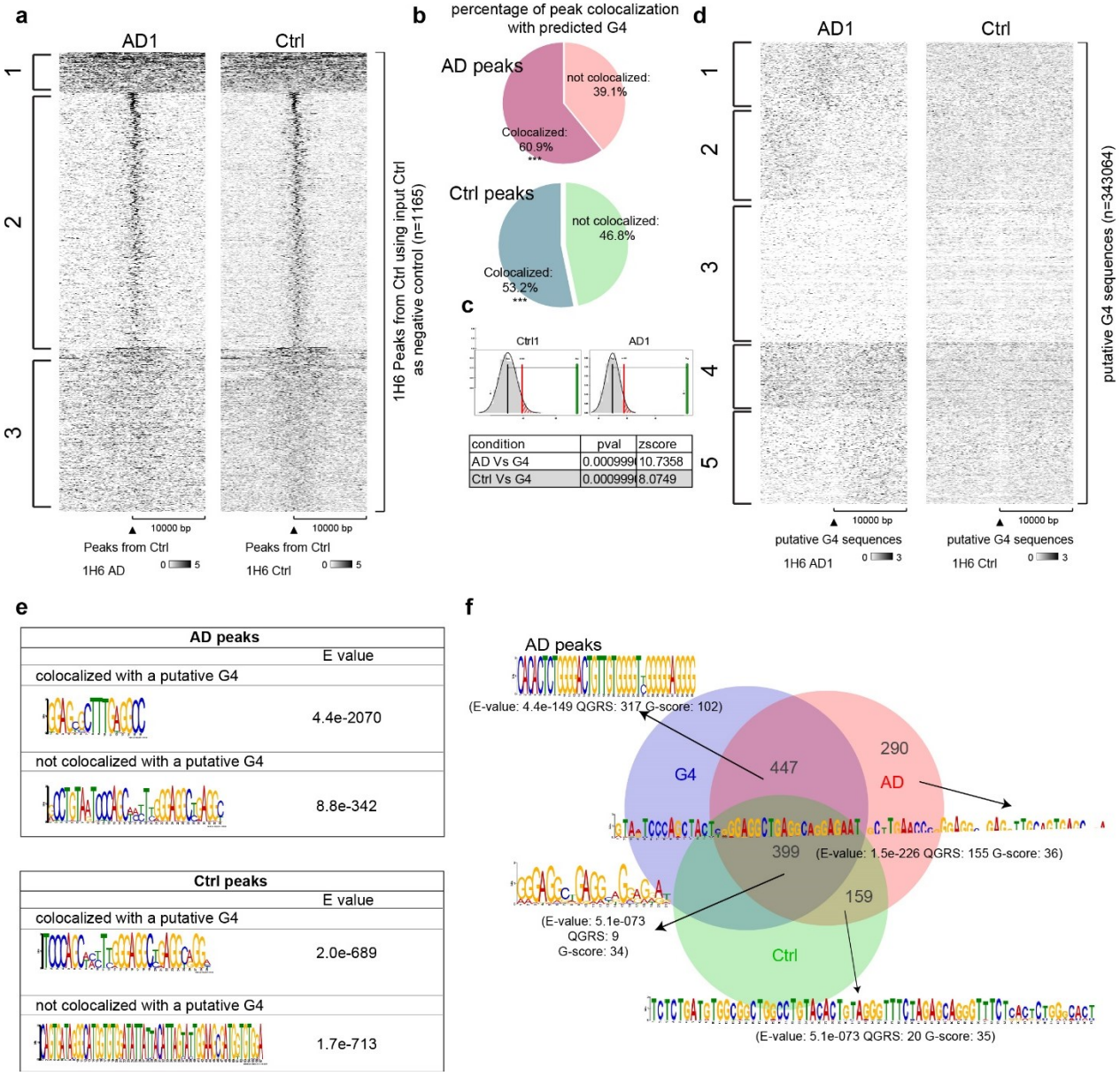


Figure 35. 1H6 peaks are more abundant in AD neurons and recognize the G4 motif.

a. Heatmap of 1H6 ChIP-seq enrichment from AD neurons or control neurons centered on the Ctrl peaks with a +/- 10 Kb. K-means clustering highlighted on the right was done using the Ctrl enrichment.

b. A pie chart showing the percentage of AD or Ctrl peaks that contain a G4 canonical sequence predicted by the Quadparser.

c. Permutation test on the colocalization between the 1H6 ChIP peaks and the G4 predicted sequences. A total of 1000 permutated sets of probes were randomly generated and annotated for G4 predicted sequences.

d. Heatmap of 1H6 ChIP-seq enrichment from AD neurons or control neurons centered on predicted G4 sequences +/- 10 Kb. K-means clustering highlighted on the right was done using the AD enrichment.

e. MEME analyses of the peaks that colocalized or not with a predicted G4 sequence, the corresponding E-value is shown for each motif.

f. Venn diagram showing the distribution of AD peaks in regard to containing or not: Ctrl peak or a predicted G4 sequence. For each category, a MEME analysis was performed, and a motif that can form a G4 structure was shown with the E-Value associated with them, the number of QGRS, and the highest G-score associated with these sequences.

3.3.3 1H6 peaks correspond to evolutionary conserved L1 sequences

Using bioinformatic analyses, we found that the best clustering approach to segregate 1H6 peaks was obtained when testing for the enrichment of repetitive elements. Indeed, upon annotation of all peaks for the presence of repetitive elements, we found that Long Interspersed Elements (*LINEs*) were present in 95% of peaks from control neurons and in 98% of peaks from AD neurons (Fig. 36a). Furthermore, 75% of the peaks also contained at least one Short Interspersed Element (SINE) (Fig. 36a). Notably, the *L1* family was the most enriched for *LINEs*, and the *ALU* family was the most enriched for SINEs (Fig. 36b-c). The enrichment for *LINEs* and *SINEs* in all 1H6 peaks was much higher than the theoretical percentage obtained when searching for *LINEs* and *SINEs* in all predicted G4 sequences of the human genome (Fig. 36a-c). These differences in enrichment could also not be accounted for by a difference in genome coverage between the mean of control and AD 1H6 peaks (~0.3%) and all predicted G4 sequences (~0.3%) (Fig. 34b). Among the *L1* family, the *LIPA3*, *LIPA2*, *LIM5*, and *LIMC4* sequences were the most represented (Fig. 36d). Among the *ALU* family, the *AluSx*, *AluY*, *AluSz*, and *AluSxz* sequences were the most represented (Fig. 37a).

Remarkably, when searching for predicted G4 motifs in the consensus sequences for these repetitions, we found that each contained at least three canonical G4 sequences (Figs. 36e and 37b).

Considering the above findings, we investigated the presence of *LINEs* or *SINEs* in intragenic 1H6 peaks. Notably, a unique peak was identified in both control and AD neurons between exons 13 and 14 of the *Amyloid beta-precursor protein* (*APP*) gene. In AD neurons, the peak was much larger than in control neurons, spanning two extra-canonical G4 sequences (Fig. 36f). We also noticed the presence of two regions with very high reads density within this unique peak (Fig. 36f-boxed peaks in red). Further analysis revealed that these regions contained four non-canonical G4 sequences presenting a relatively high G4 score (Fig. 36f). Strikingly, we found a near-perfect match between the presence of a broad intragenic 1H6 peak and the presence of a unique and evolutionary conserved *L1* sequence at the *APP* locus and all other tested loci (Figs. 36f and 37c). While *SINEs* were frequently present within or close to 1H6 peaks, they were not as predominant as *L1* sequences (Figs. 36f and 37c). Noticeably, the presence of "bystander" *SINEs* apparently accounted for the enlarged 1H6 peak found in AD neurons at the *APP* locus (Fig. 36f). These findings suggested that evolutionary conserved *L1* sequences represent the source of about all G4 structures detected in healthy and AD neurons.

Evolutionary conserved *L1* sequences are the only *LINEs* with intact internal promoters, allowing them to be transcribed by RNAPII. To test if the formation of G4 structures was dependent on transcription, we treated control and AD neurons with the RNAPII inhibitor DRB for 8 hours prior to immunofluorescence analysis (Fig. 36g). We found that inhibition of RNAPII mildly impacted G4 structures in control neurons, but dramatically reversed the accumulation of G4 structures in AD neurons (Fig. 36g). Accordingly, the 1H6 signal largely co-localized with that of RNAPII in AD neurons (Pearson correlation analysis: 0.63) (Fig. 36h). To confirm this, we treated control neurons with HDACi or HDACi + DRB. While HDACi treatment resulted in the induction of G4 structures and H3K9 acetylation, the addition of DRB prevented the induction of G4 structures independently of H3K9 acetylation (Figs. 36i and 38a). Comparable results were obtained with AD neurons (Fig. 38b). These results thus distinguished RNAPII-mediated G4 induction from chromatin relaxation.

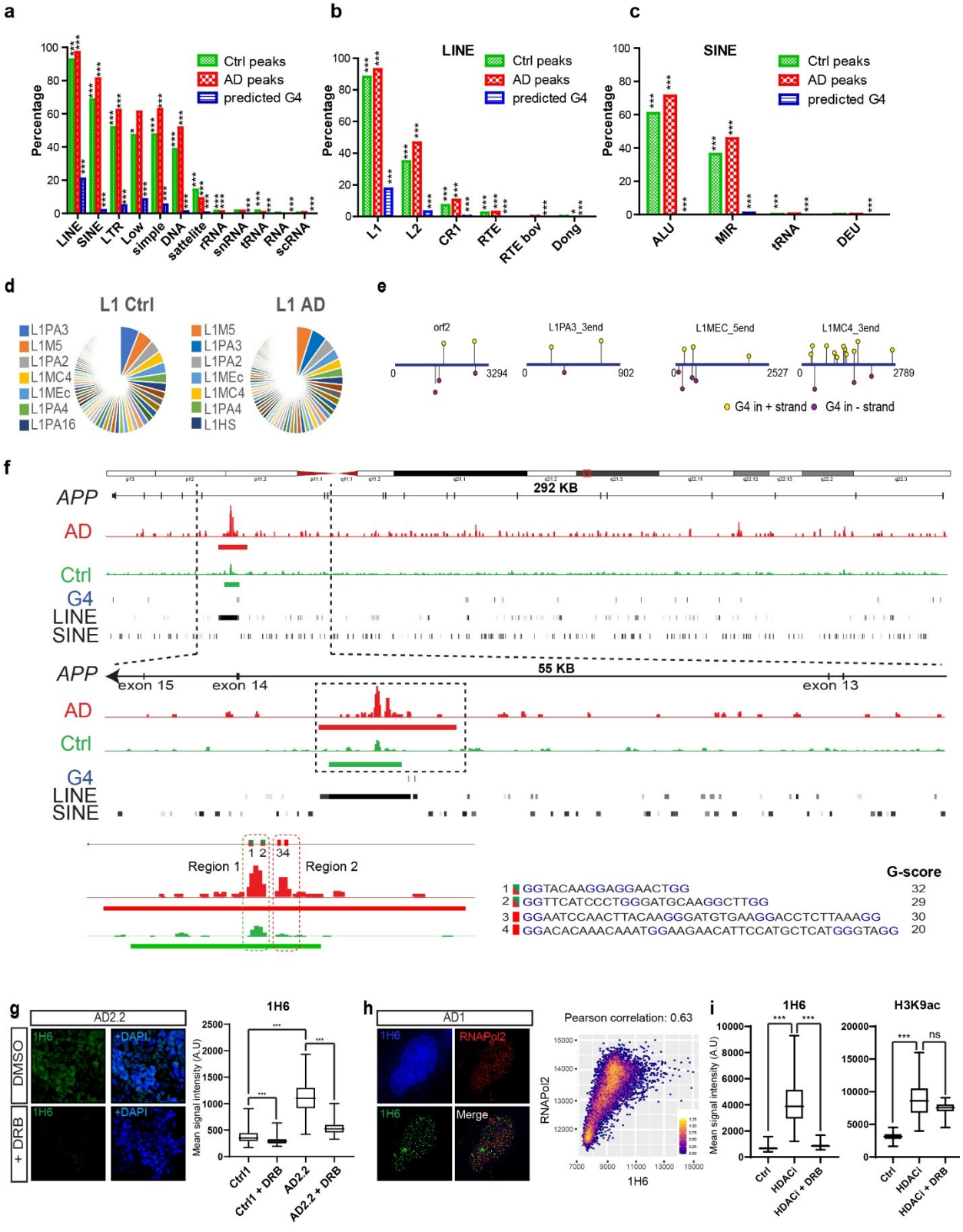


Figure 36. 1H6 peaks correspond to evolutionary conserved *L1* sequences

a. Histogram showing the distribution of 1H6 peaks and predicted G4 sequences within families of repetitive elements present in the human genome. One peak can contain a combination of various repeat elements. A permutation test was done for each combination in order to calculate the significance. A total of 1000 permutated sets of probes were randomly generated and annotated for the repeat elements.

b. Histogram showing the distribution of 1H6 peaks and predicted G4 sequences within members of the *LINE* family. One peak can contain a combination of various repeat elements. A permutation test was done for each combination in order to calculate the significance. A total of 1000 permutated sets of probes were randomly generated and annotated for the repeat elements.

c. Histogram showing the distribution of 1H6 peaks and predicted G4 sequences within members of the *SINE* family. One peak can contain a combination of various repeat elements. A permutation test was done for each combination in order to calculate the significance. A total of 1000 permutated sets of probes were randomly generated and annotated for the repeat elements.

d. Pie chart showing the repartition of repeats from *L1* sequences-the most represented *LINE* family in AD and Ctrl peaks.

e. Graphical representation of the position of all predicted G4 sequences prevalent in the *L1* family. The consensus sequence for each sub-classes of *L1* was used.

f. Physical map showing enrichment of 1H6 peaks (red: AD neurons; green: Ctrl neurons) within a unique region of the *APP* locus and spanning 6-7kb of genomic DNA. Predicted canonical G4 sequences are shown in blue. LINEs and SINEs present at the *APP* locus are represented in shades of greys. The shade of greys within LINEs or SINEs reflects their degree of conservation, with black being the most conserved. Lighter shades indicate the presence of base mismatch, base deletion, and base insertion. A zoom of the peak is represented and showing two regions of high reads and containing several G4 motifs. We used the Quadparser algorithm to measure the corresponding G-score of each motif.

g. Immunofluorescence analysis of iPSC-derived neurons treated or not with DRB for 8h, as presented in (H). Neurons were fixed with paraformaldehyde and immunolabeled with 1H6 before counterstaining with DAPI. The associated graphs show the quantification of 1H6 fluorescence intensity/cell. All values are means ± SEM. *** $P < 0.001$, Student's unpaired t-test.

h. Formaldehyde fixed AD neurons were immunolabeled with RNAPII and 1H6, and counterstained with DAPI. A zoom of a single cell is shown. Pearson correlation analysis was conducted on the AD samples. This revealed a strong correlation (0.63) between RNAPII and 1H6.

i. Mean signal intensity of 1H6 and H3K9ac in the nucleus from neurons treated with HDACi, or HDACi + DRB. Values were measured and plotted in a box and whisker graph. Statistical differences were analyzed using unpaired T-test with two tails. Scale bar: 35 µm. $P \leq 0.05$*, ≤0.01**, ≤0.001***. (see Fig. 38a)

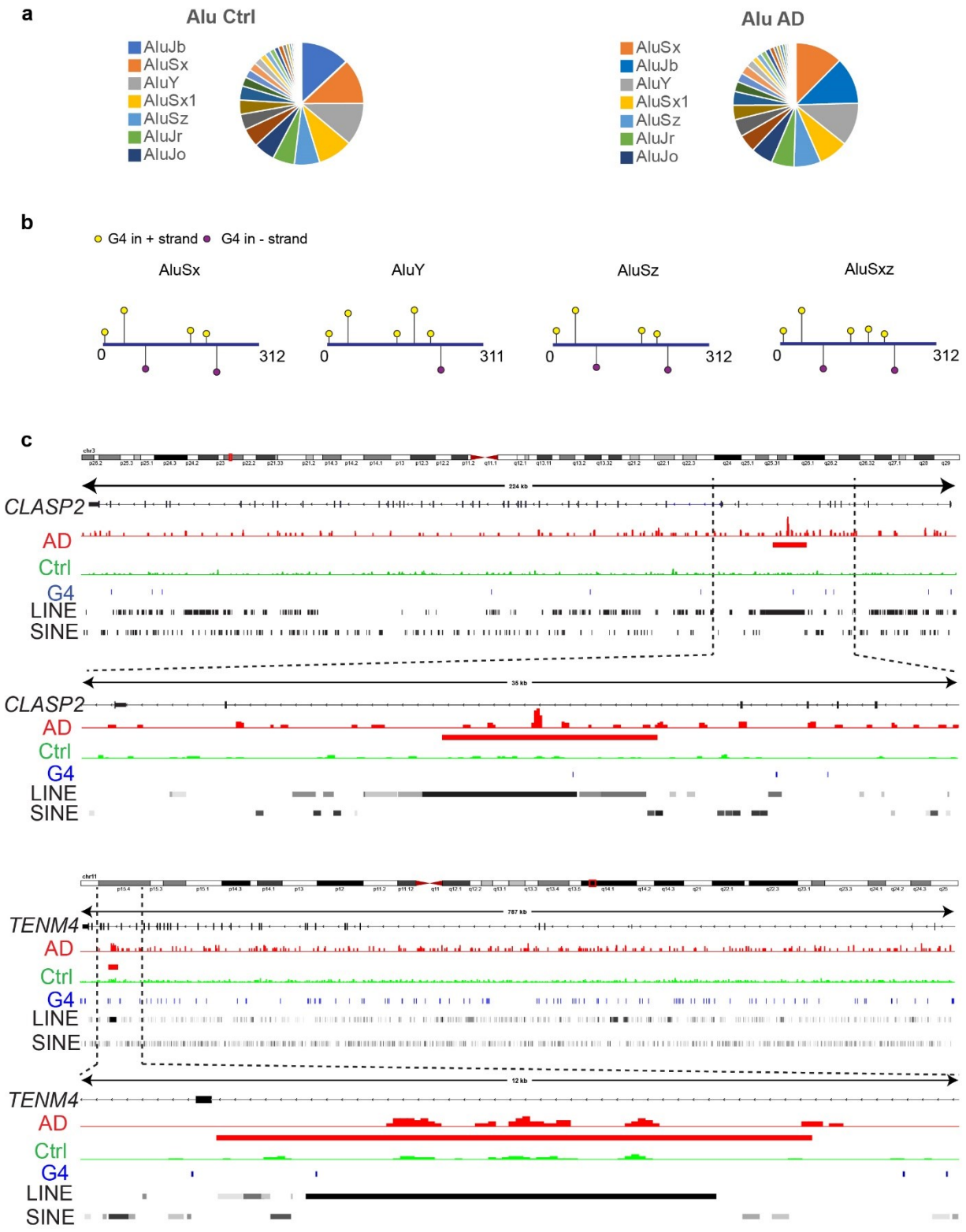


Figure 37. *SINEs* contain predicted G4 sequences

a. A pie chart showing the repartition of the repeats from *ALU*, the most represented *SINE* family in the AD and control peaks.

b. A visual representation showing the predicted G4 sequences on the consensus sequence of the repeats that are prevalent in the *ALU* family.

c. Physical maps showing enrichment of 1H6 ChIP-seq in two different gene bodies and the corresponding identified peaks. Along with the 1H6 enrichment, we plotted the repeat elements *LINE* and *SINE* with, the color of these repeats reflects their conservation with black being the most conserved and lighter shades indicate the presence of base mismatch, base deletion, and base insertion.

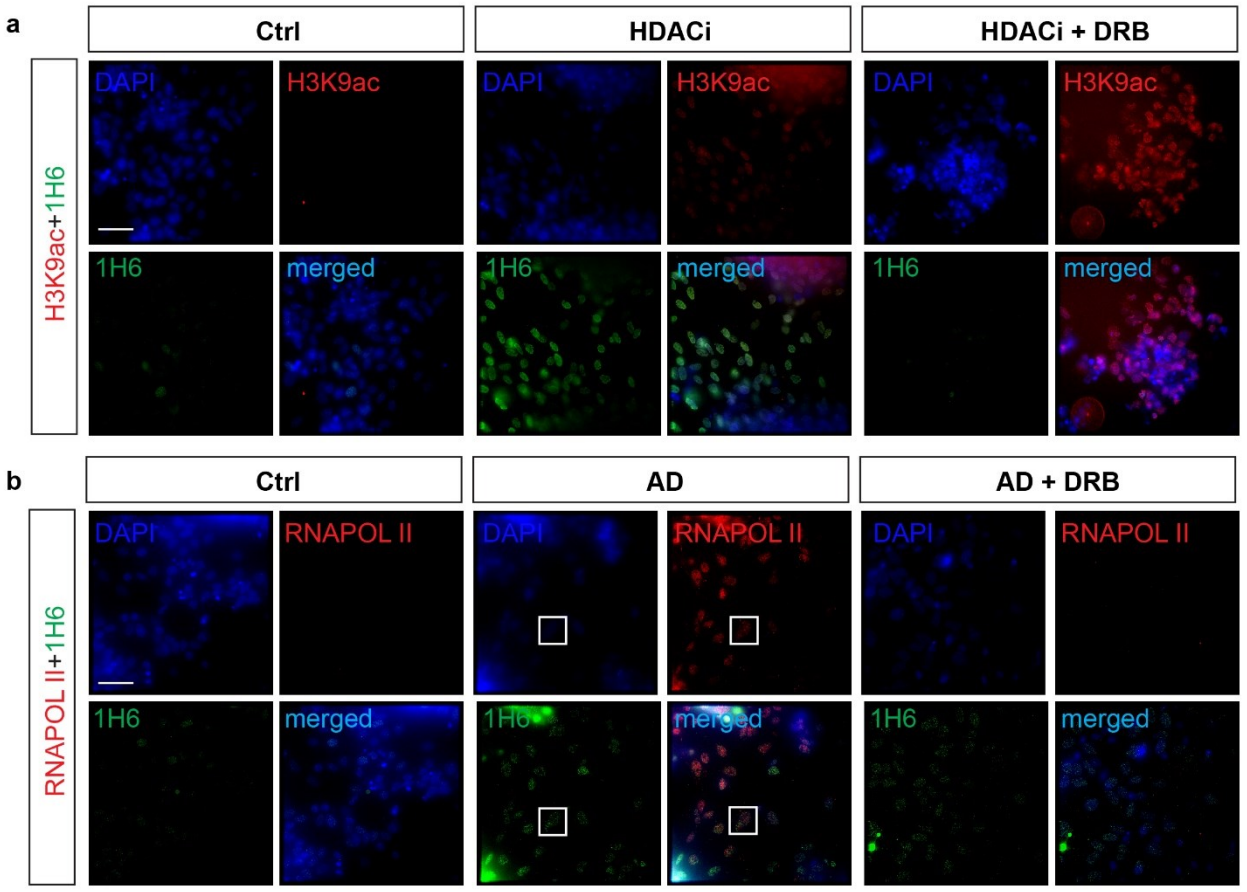


Figure 38. Transcription inhibition can rescue the G4 structures phenotype.

a. Ctrl neurons were treated with HDACi or HDACi and DRB, then fixed with formaldehyde. These cells were immunolabeled with H3K9ac and 1H6, counterstained with DAPI.

b. Ctrl or AD neurons treated for 24h with DRB then fixed with formaldehyde. These cells were immunolabeled with RNAPOL II and 1H6, counterstained with DAPI. Scale bar: 35 µm.

3.3.4 Induction of G4 structures perturb splicing and gene expression in AD neurons

To evaluate the biological significance of our findings, we performed a Gene Ontology (GO) analysis of genes associated with 1H6 peaks. Among the 139 control-specific peaks associated with a gene (i.e. lost in AD), there was no pathway enrichment. Among the 259 genes with peaks common to control and AD neurons (i.e. common peaks), there was an over-representation of genes involved in deubiquitination (Fig. 39a). Among the 269 AD-specific peaks associated with a gene (i.e. gained in AD), there was a general over-representation of genes involved in cell-cell adhesion, axonal projection, neurogenesis, and synaptogenesis (Fig. 39a). Notably, comparative RNA-seq analysis of control and AD neurons revealed that a subset of genes with intergenic or intragenic peaks presented modified splicing events (Figs. 39b, c, and 40a). In the first scenario (lost in AD), a single intergenic peak between *SMN1* and *NAIP* was lost in AD neurons. This resulted in cryptic splicing and/or transcription from *SMN1* over the adjacent *NAIP* locus, possibly explaining the presence of aberrant and larger SMN1 protein isoforms in AD neurons (Fig. 39b, d). In the second and third scenarios (common peaks and gained in AD), the presence of a single intragenic peak was found to be closely associated with alternative splicing of minor isoforms and/or exon exclusion events, possibly explaining perturbations in APP, DNAH6 and ANO4 protein isoforms observed in AD neurons (Fig. 39b, d). Efficient exon splicing can be associated with increased gene expression, while exon exclusion results in the opposite trend, a process called *exon-mediated activation of transcription starts* (759). Additionally, G4 structures located in intragenic regions have been shown to represent an obstacle to transcriptional elongation, resulting in reduced gene expression (760). To test if our results matched any of these models, we super-imposed those genes associated with aberrant splicing events over a Volcano plot

distribution of all genes differentially expressed between control and AD neurons. We found that a large proportion of genes associated with aberrant splicing events were also significantly downregulated in AD neurons (common peaks, $P = 0.0078$; gained in AD, $P = 0.002$). These results suggested that intergenic G4 structures may sometime work as “gene insulator” elements in normal conditions, and that excessive formation of intragenic G4 structures can perturb alternative splicing and gene expression in AD neurons.

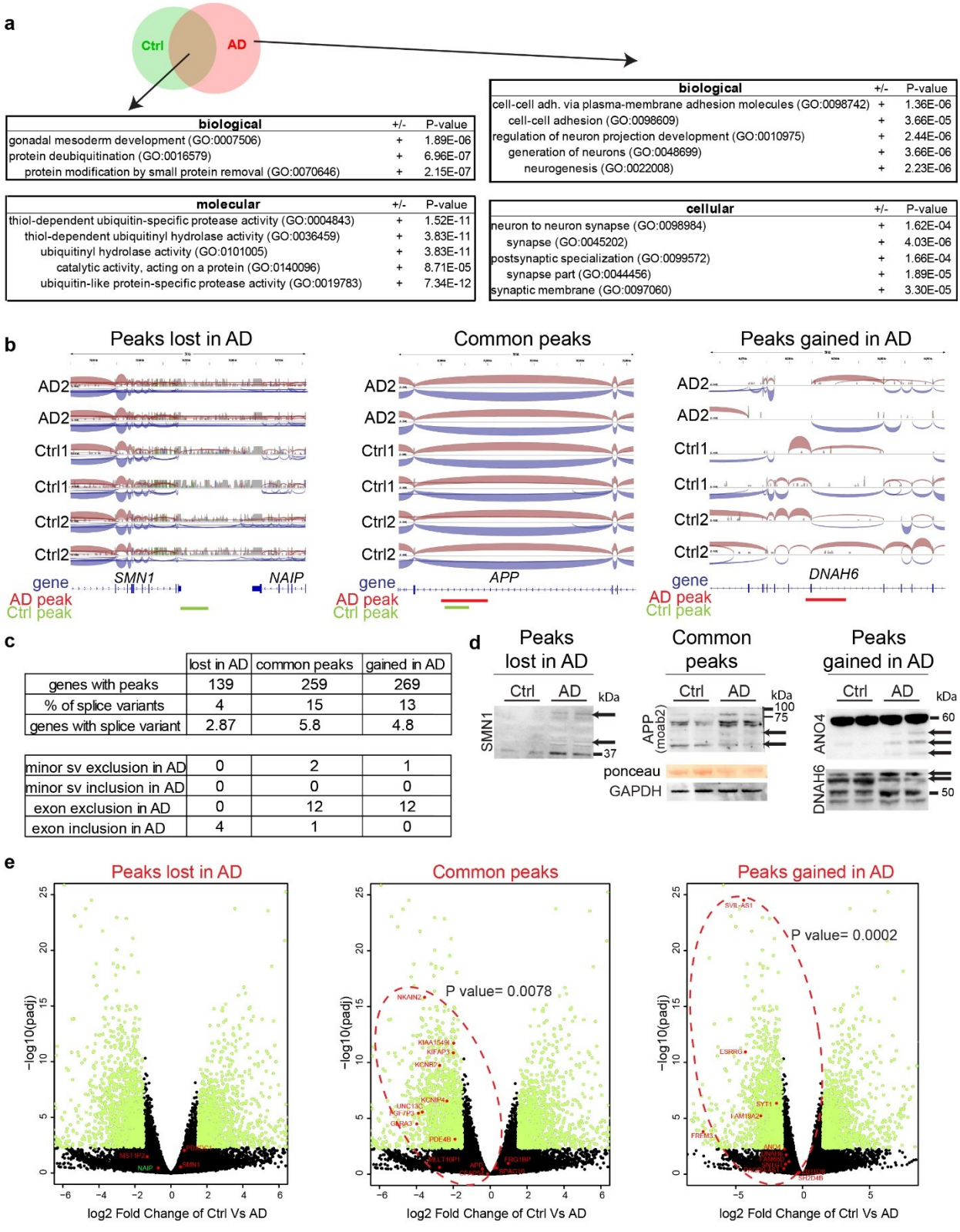


Figure 39. 1H6 peaks are enriched at key regulatory elements of neuronal genes

a. Venn diagram and gene ontology analysis of genes associated with peaks common to both AD and Ctlr neurons or specific for AD neurons (AD). The most significant pathways are indicated on the diagram along with the respective P-value.

b. Sushi plot of three different loci containing a 1H6 peak and associated with a differential splicing event. Exons are represented as blue boxes. The red lines represent a peak found in AD neurons; the green lines represent a peak found in control (Ctrl) neurons green. Loss of the intragenic peak in AD neurons between the *SMN1* and *NAIP* genes is associated with the formation of cryptic fusion transcripts. The common peak found at the *APP* locus (but increased in AD) is associated with the loss of some minor isoforms in AD neurons. The unique peak gained in AD neurons at the *DNAH6* locus is associated with loss of minor isoforms and a large splicing gap in AD neurons.

c. Table showing the number of genes containing a 1H6 peak, as well as the number of genes that entailed a differential splicing event associated with a 1H6 peak. The splicing events were subdivided into the loss of a minor isoform or formation of a gap junction.

d. Western blot analyses of Ctrl and AD neurons. GAPDH and ponceau were used as normalizers. Black arrows indicate differentially expressed protein variants.

e. Volcano plot showing differential gene expression between Ctrl vs AD neurons using RNA-seq. Superimposed in red are genes containing a G4 peak and associated with an abnormal splice event. Note that genes in the second and third volcano plot (i.e. common and gained in AD) are significantly downregulated. For each gene distribution (red dotted ovals), a P-value was calculated using the binomial distribution, knowing that 51% of the genes were downregulated.

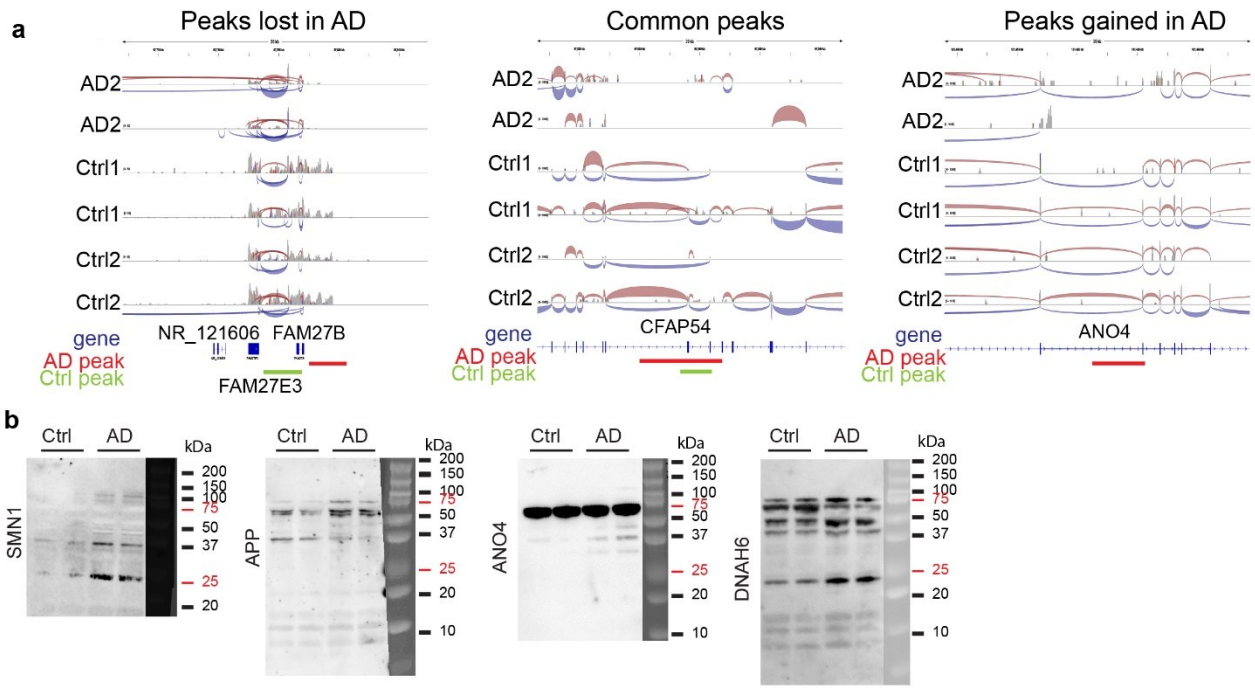


Figure 40. G4 structures affect the splicing of genes

a. Sushi plot of three different genes containing a 1H6 peak and having a differential splicing event. The blue line represents the genes with the exons, followed by the lines that represent the 1H6 peak (red: AD neurons; green: Ctrl neurons).

b. WB analysis of some of the identified proteins in Fig. 39C.

3.4 Discussion

We reported here that BMI1 inactivation in human cells or mouse photoreceptors resulted in heterochromatin relaxation and induction of G4 structures. A similar phenotype was observed in AD neurons. Loss of chromatin compaction in general, as shown using HDACi, also resulted in the induction of G4 structures. Conversely, the formation of G4 structures in AD neurons was reversed by re-establishing chromatin compaction. ChIP-seq analysis using 1H6 further revealed that most significant consensus motifs contained canonical and non-canonical G4 sequences, and that 1H6 peaks corresponded to evolutionary conserved L1 sequences. Hence, inhibition of RNAPII allowed distinguishing transcription-mediated induction of G4 structures from chromatin relaxation. 1H6 enrichment was predominant in AD neurons, and most AD-specific peaks were linked to genes involved in neurogenesis and synaptogenesis. At last, some intragenic peaks enriched in AD neurons were associated with alternative splicing events, exon exclusion, and reduced gene expression.

Herein, we have preferentially used the 1H6 antibody, which was shown to recognize several, but not all, G4 structures *in vitro* (108). We have further validated the specificity of this antibody using several approaches and by showing: 1) specific labeling after pyridostatin treatment, but not after gamma irradiation; 2) co-localization with the Werner and XPB helicases; 3) equivalent results when using the BG4 antibody; 4) specific labeling only under native chromatin conditions and; 5) detection of canonical and non-canonical G4 sequences in ChIP-seq consensus motifs. Previously, high-throughput sequencing of indirect DNA immunoprecipitation experiments with the BG4 antibody suggested that G4 DNA-forming sequences were abundant in gene bodies (112,761). ChIP-seq analyses using BG4 further confirmed that, under physiological conditions, G4 structures were enriched at the promoter of some actively transcribed genes, such as c-*MYC* (741). These observations are consistent with our findings in control neurons, where several G4 peaks were detected at intragenic and enhancer regions.

In AD neurons, heterochromatin compaction is reduced, and we found that G4 structures preferentially accumulated at LADs and peri-nucleolar heterochromatin. Likewise, heterochromatin compaction was reduced, and G4 structures were induced in *BMI1* knockdown

neurons. Notably, most *LINEs* are localized at LADs (754,762), and we indeed found that about all 1H6 peaks corresponded to *L1* sequences. In this context, it is notable that BMI1 co-purifies with LMNA and LMNB2, alongside with architectural heterochromatin proteins (731), suggesting that BMI1 may, directly or indirectly, be important for the stabilization of LADs. L1 sequences represent nearly 17% of the human genome (262). While these sequences can be autonomously transcribed by RNAPII from a 5' UTR promoter, it is estimated that out of the 500,000 L1 sequences present in the human genome, less than 5000 have an intact internal promoter (763). Notably, it was found that the 3' UTR of L1 sequences can form G4 structures *in vitro* and that this feature is conserved among mammals, suggesting a possible functional role for the propagation of retro-transposons (764,765). Strikingly, we have found that only evolutionary conserved L1 sequences i.e. predicted to be transcriptionally active, were associated with 1H6 peaks. This is consistent with our observations that the accumulation of G4 structures required functional RNAPII activity. We also showed that upon RNAPII inhibition, chromatin relaxation could be uncoupled from the accumulation of G4 structures. This supports a model where chromatin-mediated inhibition of L1 transcription is the primary mechanism regulating the formation of G4 structures in post-mitotic neurons.

A recent study revealed that many genes in AD brains displayed alternative splicing events (766), but the underlying mechanism remained unknown. However, part of these splicing anomalies were mimicked in neurons over-expressing Tau (766), thus possibly linking Tau-mediated chromatin relaxation to specific alternative splicing events. Here, we found that some intragenic 1H6 peaks enriched in AD neurons were associated with alternative splicing events and reduced gene expression, thus providing a plausible mechanism for a fraction of genes with perturbed splicing in AD. The overrepresentation of 1H6 peaks in genes involved in neurogenesis, axonal guidance, and synaptogenesis in AD samples is also intriguing considering the reported inability of AD neurons to guide their projections and form functional synapses (767–774). It is also intriguing considering that most genes with an *exon-mediated activation of transcription starts* structure are also enriched for brain development, neurogenesis, and synaptogenesis (759). Further work is clearly required to understand better the complex interconnections between G4 structures, splicing, and neuronal gene expression.

In conclusion, while specialized DNA helicases have evolved to recognize and resolve G4 structures to promote transcription, replication, and repair, chromatin-mediated transcriptional repression of L1 sequences represents a critical mechanism to prevent excessive formation of G4 structures in human neurons. Considering that heterochromatin relaxation in neurons is also observed in other neurodegenerative diseases, such as amyotrophic lateral sclerosis, frontotemporal dementia, and tauopathies (170,660,738,775–778), these findings have potentially broader implications. Our findings thus expose for the first time the nature of DNA sequences capable of forming G4 structures in human neurons, the mechanisms underlying their activation, and the biological impact of their deregulation.

3.5 Acknowledgments

This work was supported by grants from the National Science and Engineering Research Council of Canada (NSERC), Canadian Institutes of Health Research (CIHR), Maisonneuve-Rosemont Hospital Foundation, and Pierre Theroux Family Foundation. A.F., R.H., and A.B. were supported by fellowships from the Molecular Biology Program of Université de Montréal.

3.6 Data availability statement

Raw data, cell lines, and reagents are available upon request.

3.7 Code availability statement

Accession code for ChIP-seq and RNA-seq data are: GSE133113

3.8 Competing interests

A.F. and G.B. are co-founders and shareholders of StemAxon^{TM}. The corporation was, however, not involved in this study.

3.9 Author contributions

Conceived and designed: G.B. and R.H.

Performed the experiments: R.H., A.F., and A.B.

Analyzed the data: G.B., R.H., A.F., and A.B.

Wrote the paper: G.B., R.H., A.F., and A.B.

3.10 Experimental procedures

3.10.1 Human samples and animals

Human pluripotent stem cells were used in accordance with the Canadian Institute Health Research (CIHR) guidelines and approved by the "Comité de Surveillance de la Recherche sur les Cellules Souches" (CSRCS) of the CIHR and Maisonneuve-Rosemont Hospital Ethic Committee. Human brain tissues were obtained from the Douglas Hospital Brain Bank after approval by the Maisonneuve-Rosemont Hospital Ethic Committee. C57Bl/6 *Bmi1*$^{-/-}$ (The Netherlands Cancer Institute, Amsterdam) and wild-type (Charles River, St-Constant, Canada) mice were used in accordance with the Animal Care Committee of the Maisonneuve-Rosemont Hospital Research Centre (Approval ID #2014-03, #2012-09).

3.10.2 Cell cultures

Normal human diploid fibroblasts (HDFs) were purchased from the Coriell Institute. HCA2 cells were kindly provided by the laboratory of Dr. Francis Rodier. HDFs were cultured with DMEM/F12 media (Invitrogen) supplemented with 10% FBS (Invitrogen) and non-essential amino acids (Invitrogen). For the HDACi experiments, cells were treated with 5ng/ml of Trichostatin A (Sigma, T1952-200UL) or 10 mM of Sodium Butyrate (Sigma, 303410-5G), and 5µM of pyridostatin (Sigma, SML0678-5MG). For the replication and transcription arrest, we used respectively: 1µg/mL of Aphidicolin from *Nigrospora sphaerica* (Sigma, A0781-1MG) and 40µM of 5,6-Dichlorobenzimidazole 1-β-D-ribofuranoside (DRB; Sigma, D1916-10MG). A concentration of 0.2µg/mL of aphidicolin or of 0.2mM of Hydroxyurea (Sigma, H8627) was used to induce replication stress. iPSC and neuronal cultures were conducted according to the methods described in (448).

3.10.3 Differentiation of human embryonic stem cells into cortical neurons

Differentiation of iPSCs into cortical neurons was performed accordingly to Flamier *et al.* (448). Ctrl1 (fibroblasts from Coriell Institute #AG04152), Ctrl2 (fibroblasts from Coriell Institute #AG09602), AD1 (fibroblasts from Coriell Institute #AG08243) and AD2 (fibroblasts from Coriell Institute #AG08259) iPSCs were dissociated using Accutase (Innovative Cell Technology #AT-104)

and platted on growth factor reduced Matrigel (Corning #356231) in PeproGrow hES cell media (PeproTech #BM-hESC) supplemented with ROCK inhibitor (Y-27632;10µM, Cayman Chemical #10005583). Upon 70% of confluency, the media was changed to DDM supplemented with B27 (1X final), Noggin (10 ng/ml, PeproTech #120-10C) and LDN193189 (0.5µM; Sigma #SML0559). The medium was changed every day. After 16 days of differentiation, the medium was changed to DDM/B27 and replenished every day. At day 24half of the medium was changed for Neurobasal A media supplemented with B27 (1X final) and changed again every three days.

3.10.4 Immunofluorescence

Eyes were extracted and fixed by immersion over-night at 4°C in 4% paraformaldehyde (PFA)/3% sucrose in PBS, pH 7.4. Samples were washed three times in PBS, cryoprotected in PBS/30% sucrose, and frozen in CRYOMATRIX embedding medium (CEM) (Thermo Shandon, Pittsburgh, PA) or in Tissue-Tek® optimum cutting temperature (O.C.T.) compound (Sakura Finetek, USA). 5 to 12 µm thick sections were mounted on Super-Frost glass slides (Fisher Scientific) and processed for immunofluorescence. For immunofluorescence labeling, sections were incubated overnight with primary antibody solutions at 4°C in a humidified chamber. After three washes in PBS, sections were incubated with secondary antibodies for 1 h at room temperature. Slides were mounted on coverslips in DAPI-containing mounting medium (Vector Laboratories CA, H-1200). Confocal microscopy analyses were performed using 60x objectives with an IX81 confocal microscope (Olympus, Richmond Hill, Canada), and images were obtained with Fluoview software version 3.1 (Olympus). The cultured cells were fixed for 15 minutes with 4% PFA, washed three times, and then permeabilized for 10 min with 0.25 Triton (Sigma, X100-500ML), cells were then blocked in PBS/2% BSA (Sigma, A7906-100G) for an hour and incubated overnight with the primary antibody. Primary antibodies used in this study are: FITC mouse anti-TRA-1-60 (BD Pharmingen, 560380), rabbit anti-SOX2 (ab97959), goat anti-NANOG (R&D systems, af1997), goat anti-S-Opsin (1:250, Santa Cruz, sc-14363), rabbit anti-H3K9me3 (1:500, Abcam, ab8898), rabbit anti-H3K9ac (1:500, Cell Signaling, 9671S), rabbit anti-WRN (1:100, Santa Cruz, sc-5629), rabbit anti-TFIIH p80 (1:200, Santa Cruz, sc-20696) targeted against XPD, rabbit anti-TFIIH p89 (1:200, Santa Cruz, sc-293) targeted against XPB, rabbit anti-53BP1 (1:100, Novus, NB100-304), rabbit anti-H2Aub (1:200, Cell Signaling, 8240S), rabbit anti-Ki67 (1:1000, Abcam, ab15580), and mouse

BG4 (1:333, Absolute antibody, Ab00174-1.1) and 1H6 antibodies recognizing G-quadruplexes. We obtained the 1H6 antibody from The European Research Institute for the Biology of Ageing. After the primary antibodies, slides were washed three times using PBS and incubated with the secondary antibodies for 1h. The Secondary antibodies are: donkey AlexaFluor488-conjugated anti-mouse (1:1000, Life Technologies), donkey AlexaFluor488-conjugated anti-rabbit (1:1000, Life Technologies), goat AlexaFluor647-conjugated anti-mouse (1:1000, Life Technologies), goat AlexaFluor texas red-conjugated anti-rabbit (1:1000, Life Technologies). Slides were then washed three times with PBS and mounted with coverslips in DAPI-containing mounting medium (Vector Laboratories CA, H-1200).

3.10.5 Quantifications and statistical analysis

For the co-localization study, random lines were drawn on individual cells using FIJI. From these lines, we plotted the intensity profile of each marker accordingly. The data collected was plotted on horizontal graphs with each marker as a separate line for the visualization of the peaks (example figure 2D). The data was also plotted in a scatter graph, using GraphPad Prism 5, to visualize the correlation between these two markers. On these sets of pairs, a Pearson correlation was calculated to quantify the correlation. For the co-expression study intensity of the signal for different markers was measured, using a mask on DAPI to identify the nucleus, and then plotted in a scatter plot using GraphPad Prism 5, to visualize the correlation between these two markers. On these sets of pairs, a Pearson correlation was calculated to quantify the correlation. For the expression study, we quantified the mean intensity of each marker in the nucleus area using the DAPI signal to identify that area. Values were plotted with a box and whisker graph. The K-means clustering was done searching for 2 or 3 groups with 20 iterations for each run. Statistical differences were analyzed using Student's t-test for unpaired samples with $P \leq 0.05$*, ≤0.01**, ≤0.001***.

3.10.6 Western Blot

Cell extracts were homogenized in the Complete Mini Protease inhibitor cocktail solution (Roche Diagnostics), followed by sonication. Protein material was quantified using the Bradford reagent. Proteins were resolved in 1x Laemelli reducing buffer by SDS-PAGE electrophoresis and

transferred to a Nitrocellulose blotting membrane (Bio-Rad). Subsequently, membranes were blocked for 1h in 5% non-fat milk-1X TBS solution and incubated overnight with primary antibodies. The antibodies used in this study are: mouse anti SMN (Santa Cruz, SC-32313), rabbit anti-MOAB (Novus, NBP2-13075), mouse anti B4 (Santa Cruz, SC-28365), rabbit anti ANOA4 (Invitrogen, PA5-62785), and rabbit anti DNAO6 (Invitrogen, PA5-57636). Membranes were then washed 3 times in 1X TBS; 0.05% Tween solution and incubated for 1h with corresponding horseradish peroxidase-conjugated secondary antibodies. Membranes were developed using the Immobilon Western (Millipore).

3.10.7 Public ChIP-seq analysis and prediction of G-quadruplexes

ChIP-seq datasets were obtained through the GEO platform using accession numbers GSE22162 (ATRX and H3K9me3), GSE44849 (XPB and XPD), GSE76688 (BG4) and GSE38273 (BMI1). BMI1 significant peaks were extracted from Meng et al. (779). The Model-based Analysis for ChIP-Seq (MACS) was used to extract significant peaks with a P-value cutoff ≤ 0.05. Peak coordinates were mapped onto hg19 genome reference using SeqMonk v0.34.0 software (Babraham Bioinformatics). Putative G-quadruplexes were predicted using Quadparser algorithm V2 running under python v2.7.11 with indicated parameters for the number of guanines in each stack (G-groups), the number of base pairs between G-groups (loop size) and the number of time the loop and a stack was repeated after the initial stack (Repeats-1) (100). For example, to perform a Quadparser on chromosome 1, searching for a stack of 5 Gs and a loop size of 1 to 7 with 4 repeats the scrip and the parameters will be: quadparser.py -f chr1.fa -r ([gG]{5,}\w{1,7}){3,}[gG]{5,}. G-quadruplexes coordinates for each set of parameters were then mapped onto hg19 genome reference using SeqMonk software. Annotation of ChIP-seq peaks with G-quadruplexes was determined by extending them 50 base pairs on each side and counting the number of overlapping predicted G-quadruplexes. SeqMINER was used for H3K9me3 ChIP-seq enrichment heatmap and k-means clustering using default parameters.

3.10.8 ChIP-seq experiment and data analysis

The ChIP-seq experiment was done using the Diagenode kit: “iDeal ChIP-seq kit for Histones” reference number: C01010051. We followed the manufacturer protocol with a minor change:

after cross-linking, we added pyridostatin to buffer B in order to stabilize already present G4 structures. Immunoprecipitation was performed using the 1H6 antibody. The quantity and size of DNA fragments were verified on the Bioanalyzer with a “DNA 1000” chip. After precipitation, the DNA libraries were prepared using the “NEBext Fast DNA Library Prep Set for Ion Torrent” reference number E6270S. the libraries were then loaded on the “Ion Chef System” and sequenced on the Ion Torrent. Raw reads were aligned on the human genome Hg19 using the torrent platform. EaSeq software (780) (http://easeq.net) was used in order to analysis the aligned data, and to call the peaks. EaSeq was also used in order to produce the heatmaps and the train plots with a window of 10 Kb. Annotation and the statistical significance were done using the RegionR package (781) using databases downloaded from UCSC table browser (782). In order to identify the motifs, we used the MEME-ChIP from the MEME suite (783,784). In order to identify the pathways of the identified genes, we used Gene Ontology (GO) (785–787). The ChIP-seq data of the 1H6 antibody performed in this publication have been deposited in NCBI's Gene Expression Omnibus (Edgar et al., 2002) and are accessible through GEO Series accession number GSE133113**.**

3.10.9 Transcriptome analysis

Total RNA from two independent biological samples was extracted using the standard procedure of Qiagen columns and assayed for RNA integrity. cDNA was prepared according to the manufacturer’s instructions (NEB library) and sequenced using the Illumina platform. Base-calling and feature count were done using Illumina software. For differential expression analysis, Dseq2 (788) was used on R program(789).

4 Chapter 4 – Discussion

4.1 Formagenetics: a new field of molecular biology

The genetic material has always fascinated biologists at large and molecular biologists specifically. How does the living transmit that information, maintain it, replicate it, and modify it according to the environment? These are the main questions of most modern molecular biology researchers. These questions have been the source of many big discoveries over the years.

Using plants in 1866, Gregor Mendel has founded the field of modern genetics. He was the first to introduce the concept of heredity, which is now taught as the laws of Mendelian inheritance. His work on peas also introduced the idea of “invisible factors” (790), referring to what we now call genes as the factors behind the observed traits.

It was until 1953 that another major breakthrough was introduced by Watson and Crick. They have won a Nobel prize for their description of the DNA as a double helix form and the “specific pairing” of the bases (6). In this same paper, the authors have suggested a theory of the replication of the DNA using the base-pairing as a template. It is worth noting that the work of Watson and Crick was largely due to the work of Rosalind Franklin that was able to establish the X-ray diffraction pattern of the B-DNA (791). These major breakthroughs have helped ignite the research on the genetic material again, and in 1961, the genetic code was established (792). It consisted of the combinations of the bases into codon that code specifically an amino acid.

With the genetic code, a new field called Genetics was established, and the search for genes began. The most important project would be the Human Genome Project that was launched in 1990 (793,794) through 2003 (795). This project was a milestone in the field of genetics not just because it was able to sequence and catalog the genes found in the human genome but also because it established the notion that the variability in the cells cannot be solely explained by the genome sequence. However, there must be other mechanism that regulate the expression of these genes.

After the Human Project Genome, the field of epigenetics (796) came to light along with many projects, most importantly: the Human methylome project (797) and the Human Histone

Modification Database. These projects have helped to shed light on the role of epigenetics in gene regulation (798,799) and DNA repair (800).

In this thesis, and after showing the importance of the G4 structure in human cells and diseases, we suggest establishing a new field in the molecular biology word named “formagenetics” (Forma: a prefix, origin: Latin, meaning: shape): the study of the secondary structures of the DNA that stray away from the canonical Watson-Crick base pairing and the B-DNA structure of the DNA (801).

4.1.1 G4 structure as a modulator of genetic material:

In this thesis, we studied one of these structures that are highly present in the human genome is the G-quadruplex structure or what is commonly known as G4 (100). These structures consist of the stacking of multiple G-tetrad: the G-tetrad is the auto-assembly of four Guanines by a Hoogsteen hydrogen bond in a planar form, the G-tetrad is stabilized by a monovalent cation that sits in the middle of the tetrad (115). The importance of this field and these structures is growing because they are more and more identified to play essential roles in various cellular functions (802), human diseases (803), gene expression (144,178), and telomere maintenance (804).

Most of the studies in the field are centered around the proteins that can unwind these structures and the DDR response. These studies have identified many helicases that can unwind the G4: XPB and XPD helicase (136) are helicases identified to target regions with high G4 predicted sites preferentially. The BLM helicase responsible for Bloom’s syndrome was also recognized as a helicase that, with a G4 DNA as a preferred substrate (135). The WRN protein, responsible for the Werner syndrome, has also been identified as an unwinding protein of the G4 (128,805,806). The same role has also been proved for the FANCJ helicase (132,164) responsible for the Fanconia Anemia syndrome. These studies have demonstrated yet again the importance of these structures since the dysfunction of these identified helicases leads to severe disease.

While these studies and many more were essential in identifying the mechanism by which the cell reacts to these structures and eliminate them, more fundamental questions remain untouched: when and why these structures appear in the genome? And more importantly, does the cell have mechanisms by which it can protect itself from the appearance of these structures?

In the third chapter of this thesis, we covered these questions establishing a direct link between chromatin condensation and the emergence of these structures in the genome. Showing that the cell, by proper compaction of the chromatin, reduces the number of spontaneously forming G4 protecting itself against their deleterious effects, like the alternative splicing of neuronal genes seen in AD. These findings can lead to other questions equally important.

4.1.1.1 The difference between canonical and noncanonical G4 structures

While the canonical G4 structures are thought to be the unimolecular G4 structures following the predictive motif where a stretch of 4 guanines is repeated four-time separated by a loop of 3 to 7, many other non-canonical G4 structures can exist, and their function can differ from that of the canonical:

4.1.1.1.1 Unimolecular G4 structures that do not follow the consensus G4 motif

We have presented in chapter one the unimolecular G4 structure, and their canonical motif, but many G4 structures do not follow this motif (93). In our ChIP-seq data, the most represented motifs were non-canonical G4 structures formed from a stretch of three G repeated four times. This result can be an indication that at least in postmitotic human neurons, most of the G4 structures were non-canonical ones. Many chemistry papers have studied the thermodynamics properties of the canonical and non-canonical loops (807–812), these research have set the basics of these properties establishing how the loop length, the length of the G4 stretch and even the sequence underlying the G4 structure can influence how much these structures are thermodynamically stable. Unfortunately, no research has been done so far to link the various G4 motifs and their uniqueness. A systematic approach studying the various G4 motifs and their biological relevance should be made because a broad approach, like the one employed in the third chapter, can be the first step, but this step should be sharpened with time.

4.1.1.1.2 Unimolecular G4 conformation

While this thesis has studied how a cell can protect itself against G4 structures and what are the fertile terrains on which these G4 structures can develop, others have started the study on how, when these conditions are met, the DNA makes the transition between a double helix to a G4 structure. Some have suggested that a transitional phase is required in the form of a G-hairpin

(813), others have studied the actual physical, molecular, and thermodynamics mechanisms that dictate the folding process of the DNA (747,814). Depending on the folding process and the underlying G4 motif, the resulting unimolecular G4 structure can be of different conformation, as shown in the first chapter (Fig. 3) (812,815,816). All these motifs cannot be pulled down under the same umbrella, a stricter study that can differentiate between these conformations should be done. For a study like this to be possible, new antibodies should be developed. In fact, all the antibodies that are now available and recognizes the G4 structures are pan antibodies and cannot discriminate between these conformations (106,108), that is why methods that can distinguish between these various conformations is a tool that is highly needed to move forward and dissect this field. Another way to study the conformation of these structures is to modify the MDS-MaPseq (817), a method already used for RNA structures, to identify the structure linked to these G4. In fact, most of the tools that we have nowadays to study G4 structures in vivo reply heavily on antibodies. In this work, and prior of using these tools, we have gone through a round of validation detailed in chapter 3.3 in order to confirm that these antibodies are recognizing G4 structures. Nevertheless, tools that rely on antibody detection can be very unreliable, that is why the news tools that are to be developed in this field should steer clear of antibody detection and move closer to molecular probes.

4.1.1.1.3 Unimolecular Vs. polymolecular G4 structures

So far, we have discussed the unimolecular G4 structures, but as saw in (Fig. 2), G4 structures can be bimolecular and tetramolecular. These G4 structures can, in part, explain the peaks found in our ChIP-seq experiment that did not colocalize with a canonical G4 motif (many of these peaks colocalized with a non-canonical G4 motif as shown by the MEME analysis). While these G4 structures require four stretches of G to be formed, these stretches of G are not on the same DNA molecule, so not sequential on the DNA. To study these G4 structures, we have to develop a new method that combines the ChIP-seq experiment developed in chapter 3 and a C5 experiment to study the conformation and folding of the genome in a 3D model of the nucleus and then we can match the G4 peaks and the proximity of these G4 stretches in this 3D model. A model like this will require the crosslinking of the DNA like in a C5 experiment, then the immunoprecipitation of this DNA with the 1H6 antibody, once immunoprecipitated each G4 structure will be barcoded in

a manner similar to the barcode used in a single cell sequencing and then sequenced. Doing so will give us the resolution needed to determine the spatial 3D conformation of the DNA in the nucleus along with the G4 state of this DNA.

4.1.2 G4 RNA

While in this thesis, we only discussed DNA G4 or the G4 structures that can arise on the DNA, many other studies have also elucidated the existence of G4 structures in the RNA molecules (818–821). In fact, in our results we have seen elevated levels of cytoplasmic G4 structures (Fig 29 b). While these observations remain preliminary results many explanations can be given to such phenomena for example: it is known that the knockdown of BMI1 increases the transcription of the repeat elements in the genome, these repeat elements harbor many putative G4 sequences that can be attributed to the induction of G4 RNA structures. These G4 structures cannot modulate the transcription of the genes, nor can they influence the genomic material, but they can influence the translation of the RNA harboring them and, by consequence, the protein resulting from that RNA. This field of study also comes with its own tools to visualizes these structures that are often transient (104).

4.1.3 Other secondary structures of the DNA

This new field of molecular biology should not be restricted to G4 structures. In fact, as we have seen in chapter 1, DNA can adopt many secondary structures like the many forms of the double helix, the G4 structures, hairpins, and many other structures. Few works of literature have been produced about these structures, and when this literature exists, is it rarely in a human context. Formagenetics should include and favor the study of all these structures as a distinct level of regulation to be added to the genomic and epigenetic levels of regulation. The study of this field will not be very trivial since most of our tools, like qPCR, siRNA, shRNA, or Crispr-Cas9, will not be useful in this field. New tools to study the conformation and the 3D architecture of the DNA should be developed.

4.2 G4 in diseases

In chapter three of this thesis, we have seen that not only these G4 structures were more present in neurons of AD patients, but that this presence might have a direct effect on the transcription and the splicing of essential genes implicated in this disease. The enrichment of these genes is not random. We suggest that the more genes are transcribed in physiological conditions, the more they are at risk of developing G4 structures when the cells lose their chromatin condensation. In light of these findings, and knowing that these G4 structures are driven by the relaxation of the chromatin, we suggest a more in-depth study of the link between G4 structures and many human diseases associated with chromatin relaxation.

4.2.1 The link between G4 structure and aging

We have shown in this thesis that chromatin condensation is essential to keep the number of G4 structures at bay, but during aging, cells endure an epigenetic shift that reduces the chromatin condensation, this shift is well studied, and there is a plethora of literature that has analyzed this shift (552,822–827). This chromatin relaxation is due to a reduction of histone proteins and chromatin remodeler proteins like the BMI1 protein. This relaxation of the chromatin is so constant that many have studied it in order to elaborate a predictive model of age (828–833), some even called the model: the epigenetic clock of aging (834). We also proved in chapter three that BMI1 is a crucial protein in the prevention against G4 structures.

This epigenetic shift during aging, chromatin relaxation, and hypomethylation leads to higher transcription of repetitive sequences and an imbalance in gene transcription. All these deregulations, in light of chapter three of this thesis, will undoubtedly lead to more G4 structures and more deregulation of gene expression. A strategy that can reduce these G4 structures can be very effective at reducing aging and providing a better quality of life to people. These strategies can consist on over-expressing BMI1 to maintain chromatin compaction, but also overexpressing helicases in order to resolve and reduce existent G4 structures. Another strategy that is more counterintuitive is to treat the patients with G4 ligands, this will increase G4 structures in these cells. Since cancers cells have elevated levels of G4 structures to begin with, compared to normal cells, this strategy will preferentially kill cancer cells. This kind of strategies, mimic the

senotherapeutics, that are nowadays being successfully used with senescent cells to reduce aging (493,835).

4.2.2 G4 in progeria syndromes

When we knocked-down WRN protein in the human fibroblast, we had a marked increase in G4 structures. That is why we can safely theorize that progeria syndromes will have elevated levels of G4 structures. In fact, progeria syndromes represent accelerated aging due to a mutation. In these syndromes, we can always notice a relaxation of chromatin comparable to the relaxation seen with healthy aging. Furthermore, most of the progeria syndromes have a mutation in the helicases that can resolve these structures like *FANCJ*, *BLOOM*, *WRN* (131,135,156,806), or they have mutations in the lamin protein (554) reducing the anchoring capability of the heterochromatin to the LAD. Adding to this the results seen in the third chapter showing the importance of these LAD domains on G4 formation, it would suggest a heterochromatin landscape very favorable for G4 structures in progeria syndromes. These G4 structures and the genes that they affect may play a crucial role in worsening the progeria phenotype.

4.2.3 G4 in other neurodegenerative diseases

In this study, we focused on sAD to study G4 structures, but many neurodegenerative diseases are closely linked to aging, with the number one risk factor for many of them is age. Like aging and Alzheimer's disease, many of these neurodegenerative diseases have a decondensed chromatin landscape (836,837). It is the case for Huntington's disease (838–841), Parkinson's disease (which has an epigenetic drift mostly seen in the methylation of the DNA (842–844)), and amyotrophic lateral sclerosis (776,845–848). That is why many strategies to combat or even lessen the progression of these diseases are epigenetic strategies aiming at modulating the chromatin landscape (849–852). Along with these strategies, we should study the G4 structures in these neurons since the chromatin landscape is very favorable for their formation. As in Alzheimer's disease discussed in chapter three, these G4 structures can influence the transcription and the splicing of genes specific to these diseases, and targeting these G4 structures can be a more aimed way to treat these diseases.

4.3 G4 structures and the integrity of the genome

While we discussed in this thesis the transcription effect of these G4 structures, one should not overlook the replication effect and the increased damage to the DNA that these structures can induce. In fact, many have observed an increase in DNA damage upon the stabilizing of G4 structures (175,853–855), but few have studied the exact relationship between these G4 structures and the DNA damage observed.

G4 structures by themselves are a significant risk factor for DNA damage. If we observe a G4 structure (Fig. 2) we can see that one strand of the DNA is forming a bulge with the stacks of the G-quartets. This DNA bulge is recognized as DNA damage, and the DDR can be recruited in order to resolve this bulge. On the other hand, if we observe the second strand, we can notice that it is left unprotected with no complementary DNA. This form of single-stranded DNA is very prone to DNA damage like ROS species, UV, physical damage… furthermore, a single-stranded DNA can be likely to form DNA:RNA hybrids called R-loop (856,857). These R-loops are often targets of nuclease activities, increasing the risk of genome instability (858–861) and inducing double-strand breaks. While the G4 structures can by themselves increase the risk of genomic instability, they can also hinder the DDR machinery by slowing it down since these structures, even if they are not recognized by the DDR should be resolved before that the DDR machinery, or the transcription machinery, can access the DNA.

4.3.1 Targeting G4 structures as a synthetic lethality treatment for cancers

In chapter three, we have seen that the genes that are most affected by these G4 structures are those transcribed in the cell. Combining our results with the interplay between G4 structures and DNA damage, especially in cells that are actively dividing and with active transcription like cancer cells, is a very promising target for therapy (862,863). Stabilizing G4 structures in cancer cells that already harbor many genomic instabilities can be a very effective synthetic lethality approach to kill specifically these cells with minimum harm to other cells that are not dividing as much and that do not already harbor so many genomic instabilities. Many drugs can be used for that purpose, as an example, pyridostatin used in our study can be used in cancer patients. This drug stabilizes, as seen in chapter three, normally occurring G4 structures in the cells. Stable G4

structure can be read as an insult to genomic integrity by the DNA repair pathways, and DDR would be recruited to this structure and in order to resolve it (142). A cell that is lacking functional DDR may see an accumulation of these structures without being able to resolve them (864). This is the case for BRCA1 and BRCA2 deficient cells, where these cells can be targeted successfully by CX-5461, a G4 ligand (180,181). This drug is also being tested in a phase one clinical trial by Cyclene Pharmaceuticals for advanced solid tumors and lymphomas (865). On another hand, DDR can induce double-strand DNA breaks, while resolving these structures, increasing the genome instability of these cells (866,867). Since cancer cells are more actively dividing and transcribing, these cells would have more normally occurring G4 structures for the pyridostatin to stabilizes, hence creating a synthetic lethality in these cancer cells. Furthermore, many G4 stabilizing agents that have been shown to interfere with the telomerase activity can reduce the tumor size (in animals) like the case of telomestatin (868,869), RHPS4 (870,871) and triethylene tetramin (872). Finally, S2T1-6OTD, a G4 stabilizer, can inhibit the transcription of c-*MYC* and *hTERT* (873), two potent oncogenes.

5 Références bibliographiques

1. Griffith, F. The Significance of Pneumococcal Types. *J. Hyg. (Lond).* **27**, 113–159 (1928) DOI:10.1017/S0022172400031879.

2. BOIVIN, A., VENDRELY, R. & VENDRELY, C. *L'acide désoxyribonucléique du noyau cellulaire, dépositaire des caractères héréditaires; arguments d'ordre analytique*. *Comptes rendus hebdomadaires des seances de l'Academie des sciences* vol. 226 1061–3 (1948).

3. Todd, A. Synthesis in the study of nucleotides. *Science* vol. 127 787–792 (1958) DOI:10.1126/science.127.3302.787.

4. Astbury, W. T. & Bell, F. O. X-ray study of thymonucleic acid [3]. *Nature* vol. 141 747–748 (1938) DOI:10.1038/141747b0.

5. Elson, D. & Chargaff, E. On the desoxyribonucleic acid content of sea urchin gametes. *Experientia* **8**, 143–145 (1952) DOI:10.1007/BF02170221.

6. Watson, J. D. & Crick, F. H. C. C. Molecular Structure of Nucleic Acids: A Structure for Deoxyribose Nucleic Acid. *Nature* **171**, 737–738 (1953) DOI:10.1038/171737a0.

7. Crick, F. H. C. THE COMPLEMENTARY STRUCTURE OF DNA. *Proc. Natl. Acad. Sci.* **40**, 756–758 (1954) DOI:10.1073/pnas.40.8.756.

8. Dickerson, R. E. DNA structure from A to Z. *Methods Enzymol.* **211**, 67–111 (1992) DOI:10.1016/0076-6879(92)11007-6.

9. Wheeler, R. File:A-DNA, B-DNA and Z-DNA.png - Wikimedia Commons. https://commons.wikimedia.org/wiki/File:A-DNA,_B-DNA_and_Z-DNA.png (2007).

10. Wing, R. *et al.* Crystal structure analysis of a complete turn of B-DNA. *Nature* **287**, 755–758 (1980) DOI:10.1038/287755a0.

11. Peck, L. J. & Wang, J. C. Energetics of B-to-Z transition in DNA. *Proc. Natl. Acad. Sci. U. S. A.* **80**, 6206–6210 (1983) DOI:10.1073/pnas.80.20.6206.

12. Wang, A. H., Fujii, S., van Boom, J. H. & Rich, A. Molecular structure of the octamer d(G-G-C-C-G-G-C-C): modified A-DNA. *Proc. Natl. Acad. Sci. U. S. A.* **79**, 3968–3972 (1982) DOI:10.1073/pnas.79.13.3968.

13. Conner, B. N., Takano, T., Tanaka, S., Itakura, K. & Dickerson, R. E. The molecular structure of d(ICpCpGpG), a fragment of right-handed double helical A-DNA. *Nature* **295**, 294–299 (1982) DOI:10.1038/295294a0.

14. Whelan, D. R. *et al.* Detection of an *en masse* and reversible B- to A-DNA conformational transition in prokaryotes in response to desiccation. *J. R. Soc. Interface* **11**, 20140454 (2014) DOI:10.1098/rsif.2014.0454.

15. Mohr, S. C., Sokolov, N. V. H. A., Chaomei, H. & Setlow, P. Binding of small acid-soluble spore proteins from Bacillus subtilis changes the conformation of DNA from B to A. *Proc. Natl. Acad. Sci. U. S. A.* **88**, 77–81 (1991) DOI:10.1073/pnas.88.1.77.

16. Lee, K. S., Bumbaca, D., Kosman, J., Setlow, P. & Jedrzejas, M. J. Structure of a protein-DNA complex essential for DNA protection in spores of Bacillus species. *Proc. Natl. Acad. Sci.* **105**, 2806–2811 (2008) DOI:10.1073/pnas.0708244105.

17. Kejnovský, E., Nejedlý, K. & Kypr, J. Factors influencing resistance of UV-irradiated DNA to the restriction endonuclease cleavage. *Int. J. Biol. Macromol.* **34**, 213–222 (2004) DOI:10.1016/j.ijbiomac.2004.04.004.

18. Becker, M. M. & Wang, Z. Origin of ultraviolet damage in DNA. *J. Mol. Biol.* **210**, 429–438 (1989) DOI:10.1016/0022-2836(89)90120-4.

19. Dextraze, M. E., Wagner, J. R. & Hunting, D. J. 5-Bromodeoxyuridine radiosensitization: Conformation-dependent DNA damage. *Biochemistry* **46**, 9089–9097 (2007) DOI:10.1021/bi062114e.

20. Mustak, M. S. *et al.* Evidence of altered DNA integrity in the brain regions of suicidal victims of Bipolar Depression. *Indian J. Psychiatry* **52**, 220–228 (2010) DOI:10.4103/0019-5545.70974.

21. Tanaka, Y. *et al.* A'-form RNA double helix in the single crystal structure of r(UGAGCUUCGGCUC). *Nucleic Acids Res.* **27**, 949–955 (1999) DOI:10.1093/nar/27.4.949.

22. Placido, D., Brown, B. A., Lowenhaupt, K., Rich, A. & Athanasiadis, A. A Left-Handed RNA Double Helix Bound by the Zα Domain of the RNA-Editing Enzyme ADAR1. *Structure* **15**, 395–404 (2007) DOI:10.1016/j.str.2007.03.001.

23. Jacobo-Molina, A. *et al.* Crystal structure of human immunodeficiency virus type 1 reverse transcriptase complexed with double-stranded DNA at 3.0 A resolution shows bent DNA. *Proc. Natl. Acad. Sci.* **90**, 6320–6324 (1993) DOI:10.1073/pnas.90.13.6320.

24. Ebright, R. H., Ebright, Y. W. & Gunasekera, A. Consensus DNA site for the Escherichia coli catabolite gene activator protein (CAP): CAP exhibits a 450-fold higher affinity for the consensus DNA site than for the E.coli lac DNA site. *Nucleic Acids Res.* **17**, 10295–10305 (1989) DOI:10.1093/nar/17.24.10295.

25. Barber, A. M., Zhurkin, V. B. & Adhya, S. CRP-binding sites: evidence for two structural classes with 6-bp and 8-bp spacers. *Gene* **130**, 1–8 (1993) DOI:10.1016/0378-1119(93)90339-5.

26. Barber, A. M. & Zhurkin, V. B. CAP binding sites reveal pyrimidine-purine pattern characteristic of DNA bending. *J. Biomol. Struct. Dyn.* **8**, 213–232 (1990) DOI:10.1080/07391102.1990.10507803.

27. Fried, M. G., Wu, H. ming & Crothers, D. M. CAP binding to B and Z forms of DNA. *Nucleic Acids Res.* **11**, 2479–2494 (1983) DOI:10.1093/nar/11.8.2479.

28. Soo Hyun Eom, Wang, J. & Steitz, T. A. Structure of Taq polymerase with DNA at the polymerase active site. *Nature* **382**, 278–281 (1996) DOI:10.1038/382278a0.

29. Kiefer, J. R., Mao, C., Braman, J. C. & Beese, L. S. Visualizing DNA replication in a catalytically active Bacillus DNA polymerase crystal. *Nature* **391**, 304–307 (1998) DOI:10.1038/34693.

30. Timsit, Y. DNA structure and polymerase fidelity. *J. Mol. Biol.* **293**, 835–853 (1999) DOI:10.1006/jmbi.1999.3199.

31. Timsit, Y. DNA structure and polymerase fidelity: A new role for A-DNA. *J. Biomol. Struct. Dyn.* **17**, 169–176 (2000) DOI:10.1080/07391102.2000.10506617.

32. Pelletier, H., Sawaya, M. R., Kumar, A., Wilson, S. H. & Kraut, J. Structures of ternary complexes of rat DNA polymerase β, a DNA template-primer, and ddCTP. *Science (80-.).* **264**, 1891–1903 (1994) DOI:10.1126/science.7516580.

33. Pohl, F. M. & Jovin, T. M. Salt-induced co-operative conformational change of a synthetic DNA: Equilibrium and kinetic studies with poly(dG-dC). *J. Mol. Biol.* **67**, 375–396 (1972) DOI:10.1016/0022-2836(72)90457-3.

34. Wang, A. H. J. *et al.* Molecular structure of a left-Handed double helical DNA fragment at atomic resolution. *Nature* **282**, 680–686 (1979) DOI:10.1038/282680a0.

35. Arnott, S., Chandrasekaran, R., Birdsall, D. L., Leslie, A. G. W. & Ratliff, R. L. Left-handed DNA helices. *Nature* **283**, 743–745 (1980) DOI:10.1038/283743a0.

36. Lafer, E. M., Sousa, R., Ali, R., Rich, A. & Stollar, B. D. *The effect of anti-Z-DNA antibodies on the B-DNA-Z-DNA equilibrium. The Journal of biological chemistry* vol. 261 6438–43 (1986).

37. Nordheim, A. *et al.* Antibodies to left-handed Z-DNA bind to interband regions of Drosophila polytene chromosomes. *Nature* **294**, 417–422 (1981) DOI:10.1038/294417a0.

38. Nordheim, A. *et al.* Negatively supercoiled plasmids contain left-handed Z-DNA segments as detected by specific antibody binding. *Cell* **31**, 309–318 (1982) DOI:10.1016/0092-8674(82)90124-6.

39. Rahmouni, A. & Wells, R. Stabilization of Z DNA in vivo by localized supercoiling. *Science (80-.).* **246**, 358–363 (1989) DOI:10.1126/science.2678475.

40. Zhang, H., Yu, H., Ren, J. & Qu, X. Reversible B/Z-DNA transition under the low salt condition and non-B-form polydApolydT selectivity by a cubane-like europium-L-aspartic acid complex. *Biophys. J.* **90**, 3203–3207 (2006) DOI:10.1529/biophysj.105.078402.

41. Kłysik, J., Stirdivant, S. M., Larson, J. E., Hart, P. A. & Wells, R. D. Left-handed DNA in restriction fragments and a recombinant plasmid. *Nature* **290**, 672–677 (1981)

DOI:10.1038/290672a0.

42. Haniford, D. B. & Pulleyblank, D. E. Facile transition of poly[d(TG)·d(CA)] into a left-handed helix in physiological conditions. *Nature* vol. 302 632–634 (1983) DOI:10.1038/302632a0.

43. Peck, L. J., Nordheim, A., Rich, A. & Wang, J. C. Flipping of cloned d(pCpG)n.d(pCpG)n DNA sequences from right- to left-handed helical structure by salt, Co(III), or negative supercoiling. *Proc. Natl. Acad. Sci. U. S. A.* **79**, 4560–4564 (1982) DOI:10.1073/pnas.79.15.4560.

44. Ho, P. S., Ellison, M. J., Quigley, G. J. & Rich, A. A computer aided thermodynamic approach for predicting the formation of Z-DNA in naturally occurring sequences. *EMBO J.* **5**, 2737–44 (1986).

45. Ellison, M. J., Kelleher, R. J., Wang, A. H. J., Habener, J. F. & Rich, A. Sequence-dependent energetics of the B-Z transition in supercoiled DNA containing nonalternating purine-pyrimidine sequences. *Proc. Natl. Acad. Sci. U. S. A.* **82**, 8320–8324 (1985) DOI:10.1073/pnas.82.24.8320.

46. Sinden, R. R. Z-DNA - an overview | ScienceDirect Topics. https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/z-dna (1994).

47. Xu, Y., Ikeda, R. & Sugiyama, H. 8-Methylguanosine: A Powerful Z-DNA Stabilizer. *J. Am. Chem. Soc.* **125**, 13519–13524 (2003) DOI:10.1021/ja036233i.

48. Zhang, F., Huang, Q., Yan, J. & Chen, Z. Histone Acetylation Induced Transformation of B-DNA to Z-DNA in Cells Probed through FT-IR Spectroscopy. *Anal. Chem.* **88**, 4179–4182 (2016) DOI:10.1021/acs.analchem.6b00400.

49. Herbert, A. *et al.* A Z-DNA binding domain present in the human editing enzyme, double-stranded RNA adenosine deaminase. *Proc. Natl. Acad. Sci.* **94**, 8421–8426 (1997) DOI:10.1073/pnas.94.16.8421.

50. Herbert, A. G. & Rich, A. A method to identify and characterize Z-DNA binding proteins

using a linear oligodeoxynucleotide. *Nucleic Acids Res.* **21**, 2669–2672 (1993) DOI:10.1093/nar/21.11.2669.

51. Herbert, A. G., Spitzner, J. R., Lowenhaupt, K. & Rich, A. Z-DNA binding protein from chicken blood nuclei. *Proc. Natl. Acad. Sci. U. S. A.* **90**, 3339–3342 (1993) DOI:10.1073/pnas.90.8.3339.

52. Herbert, A., Lowenhaupt, K., Spitzner, J. & Rich, A. Chicken double-stranded RNA adenosine deaminase has apparent specificity for Z-DNA. *Proc. Natl. Acad. Sci. U. S. A.* **92**, 7550–7554 (1995) DOI:10.1073/pnas.92.16.7550.

53. Kim, Y.-G. *et al.* A role for Z-DNA binding in vaccinia virus pathogenesis. *Proc. Natl. Acad. Sci.* **100**, 6974–6979 (2003) DOI:10.1073/pnas.0431131100.

54. Bechert, T., Diekmann, S. & Amdtjovin, D. J. Human 170 kda and 180 kda topoisomerases ii bindpreferentially to curved and left-handed linear dna. *J. Biomol. Struct. Dyn.* **12**, 605–623 (1994) DOI:10.1080/07391102.1994.10508762.

55. Glikin, G. C., Jovin, T. M. & Arndt-jovin, D. J. Interactions of Drosophilla DNA topoisomerase II with left-handed z-DNA in supercoiled minicircles. *Nucleic Acids Res.* **19**, 7139–7144 (1991) DOI:10.1093/nar/19.25.7139.

56. Lee, A. R., Kim, N. H., Seo, Y. J., Choi, S. R. & Lee, J. H. Thermodynamic model for B-Z transition of DNA induced by Z-DNA binding proteins. *Molecules* vol. 23 (2018) DOI:10.3390/molecules23112748.

57. Wang, G. & Vasquez, K. M. Z-DNA, an active element in the genome. *Frontiers in Bioscience* vol. 12 4424–4438 (2007) DOI:10.2741/2399.

58. Liu, L. F. & Wang, J. C. Supercoiling of the DNA template during transcription. *Proc. Natl. Acad. Sci. U. S. A.* **84**, 7024–7027 (1987) DOI:10.1073/pnas.84.20.7024.

59. Schroth, G. P., Chou, P. J. & Ho, P. S. Mapping Z-DNA in the human genome. Computer-aided mapping reveals a nonrandom distribution of potential Z-DNA-forming sequences in human genes. *J. Biol. Chem.* **267**, 11846–11855 (1992).

60. Prescott, D. M. The DNA of ciliated protozoa. *Microbiol. Rev.* **58**, 233–67 (1994).

61. Lipps, H. J. *et al.* Antibodies against Z DNA react with the macronucleus but not the micronucleus of the hypotrichous ciliate stylonychia mytilus. *Cell* **32**, 435–441 (1983) DOI:10.1016/0092-8674(83)90463-4.

62. Wittig, B., Dorbic, T. & Rich, A. Transcription is associated with Z-DNA formation in metabolically active permeabilized mammalian cell nuclei. *Proc. Natl. Acad. Sci. U. S. A.* **88**, 2259–2263 (1991) DOI:10.1073/pnas.88.6.2259.

63. Wolfl, S., Martinez, C., Rich, A. & Majzoub, J. A. Transcription of the human corticotropin-releasing hormone gene in NPLC cells is correlated with Z-DNA formation. *Proc. Natl. Acad. Sci.* **93**, 3664–3668 (1996) DOI:10.1073/pnas.93.8.3664.

64. Wittig, B., Wölfl, S., Dorbic, T., Vahrson, W. & Rich, A. Transcription of human c-myc in permeabilized nuclei is associated with formation of Z-DNA in three discrete regions of the gene. *EMBO J.* **11**, 4653–63 (1992).

65. Wölfl, S., Wittig, B. & Rich, A. Identification of transcriptionally induced Z-DNA segments in the human c-myc gene. *BBA - Gene Struct. Expr.* **1264**, 294–302 (1995) DOI:10.1016/0167-4781(95)00155-7.

66. Shin, S. I. *et al.* Z-DNA-forming sites identified by ChIP-Seq are associated with actively transcribed regions in the human genome. *DNA Res.* **23**, 477–486 (2016) DOI:10.1093/dnares/dsw031.

67. Peck, L. J. & Wang, J. C. Transcriptional block caused by a negative supercoiling induced structural change in an alternating CG sequence. *Cell* **40**, 129–137 (1985) DOI:10.1016/0092-8674(85)90316-2.

68. Ditlevson, J. V. *et al.* Inhibitory effect of a short Z-DNA forming sequence on transcription elongation by T7 RNA polymerase. *Nucleic Acids Res.* **36**, 3163–3170 (2008) DOI:10.1093/nar/gkn136.

69. Aplan, P. D., Raimondi, S. C. & Kirsch, I. R. Disruption of the SCL gene by a t(1;3)

translocation in a patient with T cell acute lymphoblastic leukemia. *J. Exp. Med.* **176**, 1303–1310 (1992) DOI:10.1084/jem.176.5.1303.

70. Boehm, T. *et al.* Alternating purine-pyrimidine tracts may promote chromosomal translocations seen in a variety of human lymphoid tumours. *EMBO J.* **8**, 2621–31 (1989).

71. Satyanarayana, K. & Strominger, J. L. DNA sequences near a meiotic recombinational breakpoint within the human HLA-DQ region. *Immunogenetics* **35**, 235–40 (1992) DOI:10.1007/bf00166828.

72. Weinreb, A., Katzenberg, D. R., Gilmore, G. L. & Birshtein, B. K. Site of unequal sister chromatid exchange contains a potential Z-DNA-forming tract. *Proc. Natl. Acad. Sci. U. S. A.* **85**, 529–533 (1988) DOI:10.1073/pnas.85.2.529.

73. Blaho, J. A. & Wells, R. D. Left-handed Z-DNA and genetic recombination. *Prog. Nucleic Acid Res. Mol. Biol.* **37**, 107–26 (1989) DOI:10.1016/s0079-6603(08)60696-0.

74. van de Sande, J. H. & Jovin, T. M. Z* DNA, the left-handed helical form of poly[d(G-C)] in MgCl2-ethanol, is biologically active. *EMBO J.* **1**, 115–120 (1982) DOI:10.1002/j.1460-2075.1982.tb01133.x.

75. Wang, G., Zhao, J. & Vasquez, K. M. Detection of cis- and trans-acting factors in DNA structure-induced genetic instability using in silico and cellular approaches. *Front. Genet.* **7**, (2016) DOI:10.3389/fgene.2016.00135.

76. Wang, G., Christensen, L. A. & Vasquez, K. M. Z-DNA-forming sequences generate large-scale deletions in mammalian cells. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 2677–2682 (2006) DOI:10.1073/pnas.0511084103.

77. Lafer, E. M. *et al.* Z-DNA-specific antibodies in human systemic lupus erythematosus. *J. Clin. Invest.* **71**, 314–21 (1983) DOI:10.1172/jci110771.

78. Sibley, J. T., Lee, J. S. & Decoteau, W. E. Left-handed 'Z' DNA antibodies in rheumatoid arthritis and systemic lupus erythematosus. *J. Rheumatol.* **11**, 633–7 (1984).

79. Allinquant, B., Malfoy, B., Schuller, E. & Leng, M. Presence of Z-DNA specific antibodies in

Crohn's disease, polyradiculoneuritis and amyotrophic lateral sclerosis. *Clin. Exp. Immunol.* **58**, 29–36 (1984).

80. Rice, G. I. *et al.* Mutations in ADAR1 cause Aicardi-Goutières syndrome associated with a type i interferon signature. *Nat. Genet.* **44**, 1243–1248 (2012) DOI:10.1038/ng.2414.

81. Herbert, A. Z-DNA and Z-RNA in human disease. *Communications Biology* vol. 2 (2019) DOI:10.1038/s42003-018-0237-x.

82. Khan, N., Kolimi, N. & Rathinavelan, T. Twisting right to left: A…A mismatch in a CAG trinucleotide repeat overexpansion provokes left-handed Z-DNA conformation. *PLoS Comput. Biol.* **11**, (2015) DOI:10.1371/journal.pcbi.1004162.

83. Suram, A., Rao, J. K. S., Latha, K. S. & Viswamitra, M. A. First evidence to show the topological change of DNA from B-DNA to Z-DNA conformation in the hippocampus of Alzheimer's brain. *NeuroMolecular Med.* **2**, 289–297 (2002) DOI:10.1385/NMM:2:3:289.

84. Geng, J. & Qu, X. Recent progress report on DNA B-Z transition modulated by rare earth-amino acid complex and Alzheimer's disease amyloid beta. *Journal of Rare Earths* vol. 28 820–823 (2010) DOI:10.1016/S1002-0721(09)60232-5.

85. Edwards, S. F., Sirito, M., Krahe, R. & Sinden, R. R. A Z-DNA sequence reduces slipped-strand structure formation in the myotonic dystrophy type 2 (CCTG)·(CAGG) repeat. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 3270–3275 (2009) DOI:10.1073/pnas.0807699106.

86. Sinden, R. R., Pytlos-Sinden, M. J. & Potaman, V. N. Slipped strand DNA structures. *Frontiers in Bioscience* vol. 12 4788–4799 (2007) DOI:10.2741/2427.

87. GELLERT, M., LIPSETT, M. N. & DAVIES, D. R. Helix formation by guanylic acid. *Proc. Natl. Acad. Sci. U. S. A.* **48**, 2013–2018 (1962) DOI:10.1073/pnas.48.12.2013.

88. Sen, D. & Gilbert, W. Formation of parallel four-stranded complexes by guanine-rich motifs in DNA and its implications for meiosis. *Nature* **334**, 364–366 (1988) DOI:10.1038/334364a0.

89. Zhang, Z., Dai, J., Veliath, E., Jones, R. A. & Yang, D. Structure of a two-G-tetrad

intramolecular G-quadruplex formed by variant human telomeric sequence in K+ solution: Insights into the interconversion of human telomeric G-quadruplex structures. *Nucleic Acids Res.* **38**, 1009–1021 (2009) DOI:10.1093/nar/gkp1029.

90. Burge, S., Parkinson, G. N., Hazel, P., Todd, A. K. & Neidle, S. Quadruplex DNA: sequence, topology and structure. *Nucleic Acids Res.* **34**, 5402–15 (2006) DOI:10.1093/nar/gkl655.

91. Kim, J., Cheong, C. & Moore, P. B. Tetramerization of an RNA oligonucleotide containing a GGGG sequence. *Nature* **351**, 331–332 (1991) DOI:10.1038/351331a0.

92. File:G-quadruplex.svg - Wikipedia. https://en.wikipedia.org/wiki/File:G-quadruplex.svg#file.

93. Guédin, A., Gros, J., Alberti, P. & Mergny, J. L. How long is too long? Effects of loop size on G-quadruplex stability. *Nucleic Acids Res.* (2010) doi:10.1093/nar/gkq639 DOI:10.1093/nar/gkq639.

94. Lane, A. N., Chaires, J. B., Gray, R. D. & Trent, J. O. Stability and kinetics of G-quadruplex structures. *Nucleic Acids Research* vol. 36 5482–5515 (2008) DOI:10.1093/nar/gkn517.

95. Zhang, A. Y. Q., Bugaut, A. & Balasubramanian, S. A sequence-independent analysis of the loop length dependence of intramolecular RNA G-quadruplex stability and topology. *Biochemistry* **50**, 7251–7258 (2011) DOI:10.1021/bi200805j.

96. Hazel, P., Huppert, J., Balasubramanian, S. & Neidle, S. Loop-length-dependent folding of G-quadruplexes. *J. Am. Chem. Soc.* **126**, 16405–16415 (2004) DOI:10.1021/ja045154j.

97. Bugaut, A. & Balasubramanian, S. A sequence-independent study of the influence of short loop lengths on the stability and topology of intramolecular DNA G-quadraplexes. *Biochemistry* **47**, 689–697 (2008) DOI:10.1021/bi701873c.

98. Phan, A. T. & Patel, D. J. Two-Repeat Human Telomeric d(TAGGGTTAGGGT) Sequence Forms Interconverting Parallel and Antiparallel G-Quadruplexes in Solution: Distinct Topologies, Thermodynamic Properties, and Folding/Unfolding Kinetics. *J. Am. Chem. Soc.* **125**, 15021–15027 (2003) DOI:10.1021/ja037616j.

99. Bedrat, A., Lacroix, L. & Mergny, J. L. Re-evaluation of G-quadruplex propensity with G4Hunter. *Nucleic Acids Res.* **44**, 1746–1759 (2016) DOI:10.1093/nar/gkw006.

100. Huppert, J. L. & Balasubramanian, S. Prevalence of quadruplexes in the human genome. *Nucleic Acids Res.* **33**, 2908–16 (2005) DOI:10.1093/nar/gki609.

101. Sahakyan, A. B. *et al.* Machine learning model for sequence-driven DNA G-quadruplex formation. *Sci. Rep.* **7**, (2017) DOI:10.1038/s41598-017-14017-4.

102. Zhang, S. *et al.* Real-time monitoring of DNA G-quadruplexes in living cells with a small-molecule fluorescent probe. *Nucleic Acids Res.* **46**, 7522–7532 (2018) DOI:10.1093/nar/gky665.

103. Doria, F. *et al.* A red-NIR fluorescent dye detecting nuclear DNA G-quadruplexes: in vitro analysis and cell imaging. *Chem. Commun.* **53**, 2268–2271 (2017) DOI:10.1039/c6cc08492c.

104. Xu, S. *et al.* Directly lighting up RNA G-quadruplexes from test tubes to living human cells. *Nucleic Acids Res.* **43**, 9575–9586 (2015) DOI:10.1093/nar/gkv1040.

105. Fernando, H., Rodriguez, R. & Balasubramanian, S. Selective recognition of a DNA G-quadruplex by an engineered antibody. *Biochemistry* **47**, 9365–9371 (2008) DOI:10.1021/bi800983u.

106. Biffi, G., Tannahill, D., McCafferty, J. & Balasubramanian, S. Quantitative visualization of DNA G-quadruplex structures in human cells. *Nat. Chem.* **5**, 182–186 (2013) DOI:10.1038/nchem.1548.

107. Biffi, G., Di Antonio, M., Tannahill, D. & Balasubramanian, S. Visualization and selective chemical targeting of RNA G-quadruplex structures in the cytoplasm of human cells. *Nat. Chem.* **6**, 75–80 (2014) DOI:10.1038/nchem.1805.

108. Henderson, A. *et al.* Detection of G-quadruplex DNA in mammalian cells. *Nucleic Acids Res.* **42**, 860–9 (2014) DOI:10.1093/nar/gkt957.

109. Hoffmann, R. F. *et al.* Guanine quadruplex structures localize to heterochromatin. *Nucleic Acids Res.* (2016) doi:10.1093/nar/gkv900 DOI:10.1093/nar/gkv900.

110. Kudlicki, A. S. G-Quadruplexes Involving Both Strands of Genomic DNA Are Highly Abundant and Colocalize with Functional Sites in the Human Genome. *PLoS One* **11**, e0146174 (2016) DOI:10.1371/journal.pone.0146174.

111. Lam, E. Y. N., Beraldi, D., Tannahill, D. & Balasubramanian, S. G-quadruplex structures are stable and detectable in human genomic DNA. *Nat. Commun.* **4**, 1796 (2013) DOI:10.1038/ncomms2792.

112. Chambers, V. S. *et al.* High-throughput sequencing of DNA G-quadruplex structures in the human genome. *Nat. Biotechnol.* (2015) doi:10.1038/nbt.3295 DOI:10.1038/nbt.3295.

113. Rhodes, D. & Giraldo, R. Telomere structure and function. *Curr. Opin. Struct. Biol.* **5**, 311–22 (1995) DOI:10.1016/0959-440x(95)80092-1.

114. Sundquist, W. I. & Klug, A. Telomeric DNA dimerizes by formation of guanine tetrads between hairpin loops. *Nature* **342**, 825–829 (1989) DOI:10.1038/342825a0.

115. Williamson, J. R., Raghuraman, M. K. & Cech, T. R. Monovalent cation-induced structure of telomeric DNA: the G-quartet model. *Cell* **59**, 871–80 (1989).

116. Tang, J. *et al.* G-quadruplex preferentially forms at the very 3' end of vertebrate telomeric DNA. *Nucleic Acids Res.* **36**, 1200–1208 (2008) DOI:10.1093/nar/gkm1137.

117. Maestroni, L., Matmati, S. & Coulon, S. Solving the telomere replication problem. *Genes* vol. 8 (2017) DOI:10.3390/genes8020055.

118. Besnard, E. *et al.* Unraveling cell type-specific and reprogrammable human replication origin signatures associated with G-quadruplex consensus motifs. *Nat. Struct. Mol. Biol.* **19**, 837–844 (2012) DOI:10.1038/nsmb.2339.

119. Foulk, M. S., Urban, J. M., Casella, C. & Gerbi, S. A. Characterizing and controlling intrinsic biases of lambda exonuclease in nascent strand sequencing reveals phasing between nucleosomes and G-quadruplex motifs around a subset of human replication origins. *Genome Res.* **125**, 725–735 (2015) DOI:10.1101/gr.183848.114.

120. Huppert, J. L. & Balasubramanian, S. G-quadruplexes in promoters throughout the human

genome. *Nucleic Acids Res.* **35**, 406–413 (2007) DOI:10.1093/nar/gkl1057.

121. Eddy, J. & Maizels, N. Selection for the G4 DNA Motif at the 5' end of human genes. *Mol. Carcinog.* **48**, 319–325 (2009) DOI:10.1002/mc.20496.

122. Sun, H., Bennett, R. J. & Maizels, N. The Saccharomyces cerevisiae Sgs1 helicase efficiently unwinds G-G paired DNAs. *Nucleic Acids Res.* **27**, 1978–1984 (1999) DOI:10.1093/nar/27.9.1978.

123. Paeschke, K., Capra, J. a & Zakian, V. a. DNA replication through G-quadruplex motifs is promoted by the Saccharomyces cerevisiae Pif1 DNA helicase. *Cell* **145**, 678–91 (2011) DOI:10.1016/j.cell.2011.04.015.

124. Ribeyre, C. *et al.* The yeast Pif1 helicase prevents genomic instability caused by G-quadruplex-forming CEB1 sequences in vivo. *PLoS Genet.* **5**, (2009) DOI:10.1371/journal.pgen.1000475.

125. Sanders, C. M. Human Pif1 helicase is a G-quadruplex DNA-binding protein with G-quadruplex DNA-unwinding activity. *Biochem. J.* **430**, 119–128 (2010) DOI:10.1042/BJ20100612.

126. Law, M. J. *et al.* ATR-X syndrome protein targets tandem repeats and influences allele-specific expression in a size-dependent manner. *Cell* **143**, 367–378 (2010) DOI:10.1016/j.cell.2010.09.023.

127. Salas, T. R. *et al.* Human replication protein A unfolds telomeric G-quadruplexes. *Nucleic Acids Res.* **34**, 4857–4865 (2006) DOI:10.1093/nar/gkl564.

128. Tang, W. *et al.* The Werner syndrome RECQ helicase targets G4 DNA in human cells to modulate transcription. *Hum. Mol. Genet.* (2016) doi:10.1093/hmg/ddw079 DOI:10.1093/hmg/ddw079.

129. Huber, M. D., Duquette, M. L., Shiels, J. C. & Maizels, N. A Conserved G4 DNA Binding Domain in RecQ Family Helicases. *J. Mol. Biol.* **358**, 1071–1080 (2006) DOI:10.1016/j.jmb.2006.01.077.

130. Fry, M. & Loeb, L. A. Human Werner syndrome DNA helicase unwinds tetrahelical structures of the fragile X syndrome repeat sequence d(CGG)(n). *J. Biol. Chem.* (1999) doi:10.1074/jbc.274.18.12797 DOI:10.1074/jbc.274.18.12797.

131. London, T. B. C. *et al.* FANCJ is a structure-specific DNA helicase associated with the maintenance of genomic G/C tracts. *J. Biol. Chem.* **283**, 36132–36139 (2008) DOI:10.1074/jbc.M808152200.

132. Wu, Y., Shin-ya, K. & Brosh, R. M. FANCJ helicase defective in Fanconia anemia and breast cancer unwinds G-quadruplex DNA to defend genomic stability. *Mol. Cell. Biol.* **28**, 4116–28 (2008) DOI:10.1128/MCB.02210-07.

133. Bharti, S. K. *et al.* Specialization among iron-sulfur cluster helicases to resolve G-quadruplex DNA structures that threaten genomic stability. *J. Biol. Chem.* (2013) doi:10.1074/jbc.M113.496463 DOI:10.1074/jbc.M113.496463.

134. Bharti, S. K., Awate, S., Banerjee, T. & Brosh, R. M. Getting ready for the dance: FANCJ Irons out DNA wrinkles. *Genes* vol. 7 (2016) DOI:10.3390/genes7070031.

135. Sun, H., Karow, J. K., Hickson, I. D. & Maizels, N. The Bloom's syndrome helicase unwinds G4 DNA. *J. Biol. Chem.* **273**, 27587–92 (1998).

136. Gray, L. T., Vallur, aarthy C., Eddy, J. & Maizels, N. G quadruplexes are genomewide targets of transcriptional helicases XPB and XPD. *Nat. Chem. Biol.* **10**, 313–8 (2014) DOI:10.1038/nchembio.1475.

137. Ward, J. D., Barber, L. J., Petalcorin, M. I., Yanowitz, J. & Boulton, S. J. Replication blocking lesions present a unique substrate for homologous recombination. *EMBO J.* **26**, 3384–96 (2007) DOI:10.1038/sj.emboj.7601766.

138. Koole, W. *et al.* A polymerase theta-dependent repair pathway suppresses extensive genomic instability at endogenous G4 DNA sites. *Nat. Commun.* **5**, (2014) DOI:10.1038/ncomms4216.

139. Valton, A. L. *et al.* G4 motifs affect origin positioning and efficiency in two vertebrate

replicators. *EMBO J.* **33**, 732–746 (2014) DOI:10.1002/embj.201387506.

140. Prorok, P. *et al.* Involvement of G-quadruplex regions in mammalian replication origin activity. *Nat. Commun.* **10**, (2019) DOI:10.1038/s41467-019-11104-0.

141. Verma, A., Yadav, V. K., Basundra, R., Kumar, A. & Chowdhury, S. Evidence of genome-wide G4 DNA-mediated gene expression in human cancer cells. *Nucleic Acids Res.* **37**, 4194–4204 (2009) DOI:10.1093/nar/gkn1076.

142. Rodriguez, R. *et al.* Small-molecule-induced DNA damage identifies alternative DNA structures in human genes. *Nat. Chem. Biol.* **8**, 301–10 (2012) DOI:10.1038/nchembio.780.

143. Fernando, H. *et al.* A conserved quadruplex motif located in a transcription activation site of the human c-kit oncogene. *Biochemistry* **45**, 7854–7860 (2006) DOI:10.1021/bi0601510.

144. Siddiqui-Jain, A., Grand, C. L., Bearss, D. J. & Hurley, L. H. Direct evidence for a G-quadruplex in a promoter region and its targeting with a small molecule to repress c-MYC transcription. *Proc. Natl. Acad. Sci. U. S. A.* **99**, 11593–8 (2002) DOI:10.1073/pnas.182256799.

145. Renčiuk, D. *et al.* G-quadruplex formation in the Oct4 promoter positively regulates Oct4 expression. *Biochim. Biophys. Acta - Gene Regul. Mech.* **1860**, 175–183 (2017) DOI:10.1016/j.bbagrm.2016.11.002.

146. Cahoon, L. A. & Seifert, H. S. An alternative DNA structure is necessary for pilin antigenic variation in Neisseria gonorrhoeae. *Science* **325**, 764–7 (2009) DOI:10.1126/science.1175653.

147. Walia, R. & Chaconas, G. Suggested Role for G4 DNA in Recombinational Switching at the Antigenic Variation Locus of the Lyme Disease Spirochete. *PLoS One* **8**, (2013) DOI:10.1371/journal.pone.0057792.

148. Kuryavyi, V., Cahoon, L. A., Seifert, H. S. & Patel, D. J. RecA-binding pilE G4 sequence essential for pilin antigenic variation forms monomeric and 5’ end-stacked dimeric parallel G-quadruplexes. *Structure* **20**, 2090–2102 (2012) DOI:10.1016/j.str.2012.09.013.

149. Dempsey, L. A., Sun, H., Hanakahi, L. A. & Maizels, N. G4 DNA binding by LR1 and its

subunits, nucleolin and hnRNP D, a role for G-G pairing in immunoglobulin switch recombination. *J. Biol. Chem.* **274**, 1066–1071 (1999) DOI:10.1074/jbc.274.2.1066.

150. Kumari, S., Bugaut, A., Huppert, J. L. & Balasubramanian, S. An RNA G-quadruplex in the 5' UTR of the NRAS proto-oncogene modulates translation. *Nat. Chem. Biol.* **3**, 218–221 (2007) DOI:10.1038/nchembio864.

151. McAninch, D. S. *et al.* Fragile X mental retardation protein recognizes a G quadruplex structure within the survival motor neuron domain containing 1 mRNA 5'-UTR. *Mol. Biosyst.* **13**, 1448–1457 (2017) DOI:10.1039/c7mb00070g.

152. Werner, O. *Werner's Syndrome and Human Aging*. vol. 190 (Springer US, 1985). DOI:10.1007/978-1-4684-7853-2.

153. Werner, O. On Cataract in Conjunction with Scleroderma. in 1–14 (1985). doi:10.1007/978-1-4684-7853-2_1 DOI:10.1007/978-1-4684-7853-2_1.

154. Yu, C. E. *et al.* Positional cloning of the Werner's syndrome gene. *Science (80-.).* **272**, 258–262 (1996) DOI:10.1126/science.272.5259.258.

155. Crabbe, L., Jauch, A., Naeger, C. M., Holtgreve-Grez, H. & Karlseder, J. Telomere dysfunction as a cause of genomic instability in Werner syndrome. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 2205–2210 (2007) DOI:10.1073/pnas.0609410104.

156. Coin, F. Mutations in XPB and XPD helicases found in xeroderma pigmentosum patients impair the transcription function of TFIIH. *EMBO J.* **18**, 1357–1366 (1999) DOI:10.1093/emboj/18.5.1357.

157. Ting, A. P. L., Low, G. K. M., Gopalakrishnan, K. & Hande, M. P. Telomere attrition and genomic instability in xeroderma pigmentosum type-b deficient fibroblasts under oxidative stress. *J. Cell. Mol. Med.* **14**, 403–416 (2010) DOI:10.1111/j.1582-4934.2009.00945.x.

158. Negrini, S., Gorgoulis, V. G. & Halazonetis, T. D. Genomic instability an evolving hallmark of cancer. *Nature Reviews Molecular Cell Biology* vol. 11 220–228 (2010) DOI:10.1038/nrm2858.

159. Wei Dai, Y. Y. Genomic Instability and Cancer. *J. Carcinog. Mutagen.* **05**, (2014) DOI:10.4172/2157-2518.1000165.

160. Lopes, J. *et al.* G-quadruplex-induced instability during leading-strand replication. *EMBO J.* **30**, 4033–4046 (2011) DOI:10.1038/emboj.2011.316.

161. De, S. & Michor, F. DNA secondary structures and epigenetic determinants of cancer genome evolution. *Nat. Struct. Mol. Biol.* (2011) doi:10.1038/nsmb.2089 DOI:10.1038/nsmb.2089.

162. Aghili, L., Foo, J., DeGregori, J. & De, S. Patterns of somatically acquired amplifications and deletions in apparently normal tissues of ovarian cancer patients. *Cell Rep.* **7**, 1310–1319 (2014) DOI:10.1016/j.celrep.2014.03.071.

163. Paeschke, K. *et al.* Pif1 family helicases suppress genome instability at G-quadruplex motifs. *Nature* **497**, 458–462 (2013) DOI:10.1038/nature12149.

164. Castillo Bosch, P. *et al.* FANCJ promotes DNA synthesis through G-quadruplex structures. *EMBO J.* **33**, 2521–33 (2014) DOI:10.15252/embj.201488663.

165. Balendra, R. & Isaacs, A. M. C9orf72-mediated ALS and FTD: multiple pathways to disease. *Nature Reviews Neurology* vol. 14 544–558 (2018) DOI:10.1038/s41582-018-0047-2.

166. Zamiri, B., Reddy, K., Macgregor, R. B. & Pearson, C. E. TMPyP4 porphyrin distorts RNA G-quadruplex structures of the disease-associated r(GGGGCC)n repeat of the C9orf72 gene and blocks interaction of RNAbinding proteins. *J. Biol. Chem.* **289**, 4653–4659 (2014) DOI:10.1074/jbc.C113.502336.

167. Ivanov, P. *et al.* G-quadruplex structures contribute to the neuroprotective effects of angiogenin-induced tRNA fragments. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 18201–18206 (2014) DOI:10.1073/pnas.1407361111.

168. Yang, X.-L. tRNA-derived G-quadruplex protects motor neurons: Fig. 1. *Proc. Natl. Acad. Sci.* **111**, 18108–18109 (2014) DOI:10.1073/pnas.1420838111.

169. Ishiguro, A., Kimura, N., Watanabe, Y., Watanabe, S. & Ishihama, A. TDP-43 binds and

transports G-quadruplex-containing mRNAs into neurites for local translation. *Genes to Cells* **21**, 466–481 (2016) DOI:10.1111/gtc.12352.

170. Berson, A. *et al.* TDP-43 Promotes Neurodegeneration by Impairing Chromatin Remodeling. *Curr. Biol.* **27**, 3579-3590.e6 (2017) DOI:10.1016/j.cub.2017.10.024.

171. Mitra, J. *et al.* Motor neuron disease-associated loss of nuclear TDP-43 is linked to DNA double-strand break repair defects. *Proc. Natl. Acad. Sci. U. S. A.* **116**, 4696–4705 (2019) DOI:10.1073/pnas.1818415116.

172. Islam, M. M., Fujii, S., Sato, S., Okauchi, T. & Takenaka, S. A selective G-quadruplex DNA-stabilizing ligand based on a cyclic naphthalene diimide derivative. *Molecules* **20**, 10963–10979 (2015) DOI:10.3390/molecules200610963.

173. Asamitsu, S., Obata, S., Yu, Z., Bando, T. & Sugiyama, H. Recent progress of targeted G-quadruplex-preferred ligands toward cancer therapy. *Molecules* vol. 24 (2019) DOI:10.3390/molecules24030429.

174. Shalaby, T. *et al.* G-quadruplexes as potential therapeutic targets for embryonal tumors. *Molecules* vol. 18 12500–12537 (2013) DOI:10.3390/molecules181012500.

175. Wang, Y. *et al.* G-quadruplex DNA drives genomic instability and represents a targetable molecular abnormality in ATRX-deficient malignant glioma. *Nat. Commun.* **10**, 943 (2019) DOI:10.1038/s41467-019-08905-8.

176. McLuckie, K. I. E. *et al.* G-quadruplex DNA as a molecular target for induced synthetic lethality in cancer cells. *J. Am. Chem. Soc.* (2013) doi:10.1021/ja404868t DOI:10.1021/ja404868t.

177. Merle, P. *et al.* Highly efficient radiosensitization of human glioblastoma and lung cancer cells by a G-quadruplex DNA binding compound. *Sci. Rep.* **5**, 16255 (2015) DOI:10.1038/srep16255.

178. Zizza, P. *et al.* Intragenic G-quadruplex structure formed in the human CD133 and its biological and translational relevance. *Nucleic Acids Res.* **44**, 1579–90 (2016)

DOI:10.1093/nar/gkv1122.

179. Balasubramanian, S., Hurley, L. H. & Neidle, S. Targeting G-quadruplexes in gene promoters: A novel anticancer strategy? *Nat. Rev. Drug Discov.* **10**, 261–275 (2011) DOI:10.1038/nrd3428.

180. Zimmer, J. *et al.* Targeting BRCA1 and BRCA2 Deficiencies with G-Quadruplex-Interacting Compounds. *Mol. Cell* **61**, 449–460 (2016) DOI:10.1016/j.molcel.2015.12.004.

181. Xu, H. *et al.* CX-5461 is a DNA G-quadruplex stabilizer with selective lethality in BRCA1/2 deficient tumours. *Nat. Commun.* **8**, (2017) DOI:10.1038/ncomms14432.

182. Genome Reference Consortium. GRCh38.p13 - Genome - Assembly - NCBI. https://www.ncbi.nlm.nih.gov/assembly/GCF_000001405.39/ (2019).

183. Alberts, B. . J. A. . L. J. . R. M. . R. K. . and W. P. Molecular biology of the cell (4th ed.). *Biochem. Mol. Biol. Educ.* **31**, 212–214 (2003) DOI:10.1002/bmb.2003.494031049999.

184. Hardy, P. A. & Zacharias, H. Walther Flemming on histology in medicine 1878: a newly discovered letter to his father. *Ann. Anat.* **191**, 171–85 (2009) DOI:10.1016/j.aanat.2009.01.002.

185. Stedman, E. & Stedman, E. The basic proteins of cell nuclei. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **235**, 565–595 (1951) DOI:10.1098/rstb.1951.0008.

186. Doenecke, D. & Karlson, P. Albrecht Kossel and the discovery of histones. *Trends Biochem. Sci.* **9**, 404–405 (1984) DOI:10.1016/0968-0004(84)90226-3.

187. Luger, K., Mäder, A. W., Richmond, R. K., Sargent, D. F. & Richmond, T. J. Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature* **389**, 251–260 (1997) DOI:10.1038/38444.

188. Olins, A. L. & Olins, D. E. Spheroid chromatin units (v bodies). *Science (80-.).* **183**, 330–332 (1974) DOI:10.1126/science.183.4122.330.

189. Davies, H. G. & Small, J. V. Structural units in chromatin and their orientation on

membranes. *Nature* **217**, 1122–1125 (1968) DOI:10.1038/2171122a0.

190. Ris, H. & Kubai, D. F. Chromosome Structure. *Annu. Rev. Genet.* **4**, 263–294 (1970) DOI:10.1146/annurev.ge.04.120170.001403.

191. Thomas, J. O. Histone H1: Location and role. *Current Opinion in Cell Biology* vol. 11 312–317 (1999) DOI:10.1016/S0955-0674(99)80042-8.

192. Woodcock, C. L. & Dimitrov, S. Higher-order structure of chromatin and chromosomes. *Current Opinion in Genetics and Development* vol. 11 130–135 (2001) DOI:10.1016/S0959-437X(00)00169-6.

193. Annunziato, A. T. DNA Packaging: Nucleosomes and Chromatin | Learn Science at Scitable. *Nat. Educ.* **1**, 1(1):26 (2008).

194. Woodcock, C. File:Chromatin nucleofilaments (detail).png - Wikimedia Commons. *Wikimedia Commons* https://commons.wikimedia.org/wiki/File:Chromatin_nucleofilaments_(detail).png (2011).

195. Wheeler, R. File:Chromatin Structures.png - Wikimedia Commons. *Wikimedia Commons* https://commons.wikimedia.org/wiki/File:Chromatin_Structures.png (2005).

196. Grosberg, A. Y., Nguyen, T. T. & Shklovskii, B. I. Colloquium: The physics of charge inversion in chemical and biological systems. *Reviews of Modern Physics* vol. 74 329–345 (2002) DOI:10.1103/RevModPhys.74.329.

197. Perico, A., La Penna, G. & Arcesi, L. Electrostatic interactions with histone tails may bend linker DNA in chromatin. *Biopolymers* vol. 81 20–28 (2006) DOI:10.1002/bip.20371.

198. Korolev, N., Lyubartsev, A. P. & Nordenskiöld, L. Computer modeling demonstrates that electrostatic attraction of nucleosomal DNA is mediated by histone tails. *Biophys. J.* **90**, 4305–4316 (2006) DOI:10.1529/biophysj.105.080226.

199. Luger, K. & Richmond, T. J. The histone tails of the nucleosome. *Curr. Opin. Genet. Dev.* **8**, 140–146 (1998) DOI:10.1016/S0959-437X(98)80134-2.

200. Arya, G. & Schlick, T. Role of histone tails in chromatin folding revealed by a mesoscopic oligonucleosome model. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 16236–16241 (2006) DOI:10.1073/pnas.0604817103.

201. Heitz, E. Das heterochromatin der moose. (1928).

202. Brown, S. W. Heterochromatin. *Science (80-.).* **151**, 417–425 (1966) DOI:10.1126/science.151.3709.417.

203. Russell, L. B. SECTION OF BIOLOGICAL AND MEDICAL SCIENCES: ANOTHER LOOK AT THE SINGLE-ACTIVE-X HYPOTHESIS*. *Trans. N. Y. Acad. Sci.* **26**, 726–736 (1964) DOI:10.1111/j.2164-0947.1964.tb01939.x.

204. Sidhu, S. K., Minks, J., Chang, S. C., Cotton, A. M. & Brown, C. J. X chromosome inactivation: heterogeneity of heterochromatin. *Biochem. Cell Biol.* **86**, 370–9 (2008) DOI:10.1139/o08-100.

205. ALLFREY, V. G., FAULKNER, R. & MIRSKY, A. E. ACETYLATION AND METHYLATION OF HISTONES AND THEIR POSSIBLE ROLE IN THE REGULATION OF RNA SYNTHESIS. *Proc. Natl. Acad. Sci. U. S. A.* **51**, 786–94 (1964).

206. Levinger, L. & Varshavsky, A. Selective arrangement of ubiquitinated and D1 protein-containing nucleosomes within the drosophila genome. *Cell* **28**, 375–385 (1982) DOI:10.1016/0092-8674(82)90355-5.

207. Gutierrez, R. M. & Hnilica, L. S. Tissue specificity of histone phosphorylation. *Science (80-.).* **157**, 1324–1325 (1967) DOI:10.1126/science.157.3794.1324.

208. Stevely, W. S. & Stocken, L. A. Phosphorylation of Rat-Thymus Histone. *Biochem. J.* **100**, 20C-21C (1966) DOI:10.1042/bj1000020c.

209. RICKWOOD, D., MacGILLIVRAY, A. J. & WHISH, W. J. D. The Modification of Nuclear Proteins by ADP-ribosylation. *Eur. J. Biochem.* **79**, 589–598 (1977) DOI:10.1111/j.1432-1033.1977.tb11843.x.

210. Shiio, Y. & Eisenman, R. N. Histone sumoylation is associated with transcriptional

repression. *Proc. Natl. Acad. Sci. U. S. A.* **100**, 13225–13230 (2003) DOI:10.1073/pnas.1735528100.

211. Waddington, C. H. The epigenotype. 1942. *Int. J. Epidemiol.* **41**, 10–13 (2012) DOI:10.1093/ije/dyr184.

212. Wu, C. T. & Morris, J. R. Genes, genetics, and epigenetics: A correspondence. *Science* vol. 293 1103–1105 (2001) DOI:10.1126/science.293.5532.1103.

213. Dupont, C., Armant, D. R. & Brenner, C. A. Epigenetics: Definition, mechanisms and clinical perspective. *Seminars in Reproductive Medicine* vol. 27 351–357 (2009) DOI:10.1055/s-0029-1237423.

214. Wei, J. W., Huang, K., Yang, C. & Kang, C. S. Non-coding RNAs as regulators in epigenetics (Review). *Oncol. Rep.* **37**, 3–9 (2017) DOI:10.3892/or.2016.5236.

215. Shirakawa, S. & Saunders, G. F. In Vivo Methylation of DNA in Human Leukocytes. *Proc. Soc. Exp. Biol. Med.* **138**, 369–372 (1971) DOI:10.3181/00379727-138-35899.

216. Desai, L. S., Wulff, U. C. & Foley, G. E. Human leukemic cells. Abnormal amount of methylated base in DNA. *Exp. Cell Res.* **65**, 260–263 (1971) DOI:10.1016/S0014-4827(71)80079-4.

217. Xiao, C.-L., Zhu, S., Gu, X.-F., Wang, K. & Yan, G.-R. N 6-Methyladenine DNA Modification in the Human Genome Graphical Abstract Highlights d DNA 6mA modification is extensively present in the human genome d Methyltransferase N6AMT1 is responsible for DNA 6mA methylation modification d Demethylase ALKBH1 is responsible for DNA 6mA demethylation modification d Decrease of genomic DNA 6mA promotes human tumorigenesis. *Mol. Cell* **71**, 306–318 (2018) DOI:10.1016/j.molcel.2018.06.015.

218. Li, E., Bestor, T. H. & Jaenisch, R. Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. *Cell* **69**, 915–926 (1992) DOI:10.1016/0092-8674(92)90611-F.

219. Li, E., Beard, C. & Jaenisch, R. Role for DNA methylation in genomic imprinting. *Nature* **366**,

362–365 (1993) DOI:10.1038/366362a0.

220. Bestor, T. H. Activation of mammalian DNA methyltransferase by cleavage of a Zn binding regulatory domain. *EMBO J.* **11**, 2611–2617 (1992) DOI:10.1002/j.1460-2075.1992.tb05326.x.

221. Pradhan, S., Bacolla, A., Wells, R. D. & Roberts, R. J. Recombinant human DNA (cytosine-5) methyltransferase. I. Expression, purification, and comparison of novo and maintenance methylation. *J. Biol. Chem.* **274**, 33002–33010 (1999) DOI:10.1074/jbc.274.46.33002.

222. Robert, M. F. *et al.* DNMT1 is required to maintain CpG methylation and aberrant gene silencing in human cancer cells. *Nat. Genet.* **33**, 61–65 (2003) DOI:10.1038/ng1068.

223. Goyal, R., Reinhardt, R. & Jeltsch, A. Accuracy of DNA methylation pattern preservation by the Dnmt1 methyltransferase. *Nucleic Acids Res.* **34**, 1182–1188 (2006) DOI:10.1093/nar/gkl002.

224. Lei, H. *et al.* De novo DNA cytosine methyltransferase activities in mouse embryonic stem cells. *Development* (1996).

225. Okano, M., Bell, D. W., Haber, D. A. & Li, E. DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. *Cell* **99**, 247–257 (1999) DOI:10.1016/S0092-8674(00)81656-6.

226. Hata, K., Okano, M., Lei, H. & Li, E. Dnmt3L cooperates with the Dnmt3 family of de novo DNA methyltransferases to establish maternal imprints in mice. *Development* (2002).

227. Kaneda, M. *et al.* Essential role for de novo DNA methyltransferase Dnmt3a in paternal and maternal imprinting. *Nature* (2004) doi:10.1038/nature02633 DOI:10.1038/nature02633.

228. Caiafa, P. & Zampieri, M. DNA methylation and chromatin structure: The puzzling CpG islands. *Journal of Cellular Biochemistry* vol. 94 257–265 (2005) DOI:10.1002/jcb.20325.

229. Siegfried, Z. *et al.* DNA methylation represses transcription in vivo. *Nat. Genet.* **22**, 203–206 (1999) DOI:10.1038/9727.

230. Rosenfeld, J. A. *et al.* Determination of enriched histone modifications in non-genic portions of the human genome. *BMC Genomics* **10**, (2009) DOI:10.1186/1471-2164-10-143.

231. Guillemette, B. *et al.* H3 lysine 4 is acetylated at active gene promoters and is regulated by H3 lysine 4 methylation. *PLoS Genet.* **7**, (2011) DOI:10.1371/journal.pgen.1001354.

232. Creyghton, M. P. *et al.* Histone H3K27ac separates active from poised enhancers and predicts developmental state. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 21931–21936 (2010) DOI:10.1073/pnas.1016071107.

233. Mahrez, W. *et al.* H3K36ac is an evolutionary conserved plant histone modification that marks active genes. *Plant Physiol.* **170**, 1566–1577 (2016) DOI:10.1104/pp.15.01744.

234. Xu, F., Zhang, K. & Grunstein, M. Acetylation in histone H3 globular domain regulates gene expression in yeast. *Cell* **121**, 375–385 (2005) DOI:10.1016/j.cell.2005.03.011.

235. Topal, S., Vasseur, P., Radman-Livaja, M. & Peterson, C. L. Distinct transcriptional roles for Histone H3-K56 acetylation during the cell cycle in Yeast. *Nat. Commun.* **10**, (2019) DOI:10.1038/s41467-019-12400-5.

236. Zhao, R., Nakamura, T., Fu, Y., Lazar, Z. & Spector, D. L. Gene bookmarking accelerates the kinetics of post-mitotic transcriptional re-activation. *Nat. Cell Biol.* **13**, 1295–1304 (2011) DOI:10.1038/ncb2341.

237. Park, C. S., Rehrauer, H. & Mansuy, I. M. Genome-wide analysis of H4K5 acetylation associated with fear memory in mice. *BMC Genomics* **14**, (2013) DOI:10.1186/1471-2164-14-539.

238. Gupta, A. P. *et al.* Histone 4 lysine 8 acetylation regulates proliferation and host-pathogen interaction in Plasmodium falciparum. *Epigenetics and Chromatin* **10**, 40 (2017) DOI:10.1186/s13072-017-0147-z.

239. Wang, Z. *et al.* Combinatorial patterns of histone acetylations and methylations in the human genome. *Nat. Genet.* **40**, 897–903 (2008) DOI:10.1038/ng.154.

240. Nagarajan, S., Benito, E., Fischer, A. & Johnsen, S. A. H4K12ac is regulated by estrogen receptor-alpha and is associated with BRD4 function and inducible transcription. *Oncotarget* **6**, 7305–7317 (2015) DOI:10.18632/oncotarget.3439.

241. Taylor, G. C. A., Eskeland, R., Hekimoglu-Balkan, B., Pradeepa, M. M. & Bickmore, W. A. H4K16 acetylation marks active genes and enhancers of embryonic stem cells, but does not alter chromatin compaction. *Genome Res.* **23**, 2053–2065 (2013) DOI:10.1101/gr.155028.113.

242. Rada-Iglesias, A. Is H3K4me1 at enhancers correlative or causative? *Nature Genetics* vol. 50 4–5 (2018) DOI:10.1038/s41588-017-0018-3.

243. Steger, D. J. *et al.* DOT1L/KMT4 Recruitment and H3K79 Methylation Are Ubiquitously Coupled with Gene Transcription in Mammalian Cells. *Mol. Cell. Biol.* **28**, 2825–2839 (2008) DOI:10.1128/mcb.02076-07.

244. Metzger, E. *et al.* KMT9 monomethylates histone H4 lysine 12 and controls proliferation of prostate cancer cells. *Nat. Struct. Mol. Biol.* **26**, 361–371 (2019) DOI:10.1038/s41594-019-0219-9.

245. Liu, Y. *et al.* H3K4me2 functions as a repressive epigenetic mark in plants. *Epigenetics and Chromatin* **12**, (2019) DOI:10.1186/s13072-019-0285-6.

246. Huang, X. *et al.* Stable H3K4me3 is associated with transcription initiation during early embryo development. *Bioinformatics* **35**, 3931–3936 (2019) DOI:10.1093/bioinformatics/btz173.

247. Zhao, B. *et al.* H3K14me3 genomic distributions and its regulation by KDM4 family demethylases. *Cell Research* vol. 28 1118–1120 (2018) DOI:10.1038/s41422-018-0095-6.

248. Mishra, M. K. & Bishnupuri, K. S. *Epigenetic advancements in cancer*. *Springer* (Springer International Publishing, 2016). doi:10.1007/978-3-319-24951-3 DOI:10.1007/978-3-319-24951-3.

249. Pertea, M. The Human Transcriptome: An Unfinished Story. *Genes (Basel).* **3**, 344–360

(2012) DOI:10.3390/genes3030344.

250. Smith, C. L. & Peterson, C. L. ATP-Dependent Chromatin Remodeling. *Curr. Top. Dev. Biol.* **65**, 115–148 (2004) DOI:10.1016/S0070-2153(04)65004-6.

251. Arney, K. L. & Fisher, A. G. Epigenetic aspects of differentiation. *J. Cell Sci.* **117**, 4355–4363 (2004) DOI:10.1242/jcs.01390.

252. Fisher, C. L. & Fisher, A. G. Chromatin states in pluripotent, differentiated, and reprogrammed cells. *Current Opinion in Genetics and Development* vol. 21 140–146 (2011) DOI:10.1016/j.gde.2011.01.015.

253. Mullen, A. C. *et al.* Cell cycle controlling the silencing and functioning of mammalian activators. *Curr. Biol.* **11**, 1695–9 (2001) DOI:10.1016/s0960-9822(01)00533-4.

254. Lu, B. Y., Ma, J. & Eissenberg, J. C. Developmental regulation of heterochromatin-mediated gene silencing in Drosophila. *Development* **125**, (1998).

255. Bernstein, B. E. *et al.* A Bivalent Chromatin Structure Marks Key Developmental Genes in Embryonic Stem Cells. *Cell* **125**, 315–326 (2006) DOI:10.1016/j.cell.2006.02.041.

256. Ohno, S. So much 'junk' DNA in our genome. *Brookhaven Symp. Biol.* **23**, 366–70 (1972).

257. de Koning, A. P. J., Gu, W., Castoe, T. A., Batzer, M. A. & Pollock, D. D. Repetitive elements may comprise over Two-Thirds of the human genome. *PLoS Genet.* **7**, e1002384 (2011) DOI:10.1371/journal.pgen.1002384.

258. Naughton, C., Sproul, D., Hamilton, C. & Gilbert, N. Analysis of Active and Inactive X Chromosome Architecture Reveals the Independent Organization of 30 nm and Large-Scale Chromatin Structures. *Mol. Cell* **40**, 397–409 (2010) DOI:10.1016/j.molcel.2010.10.013.

259. Cordaux, R. & Batzer, M. A. The impact of retrotransposons on human genome evolution. *Nature Reviews Genetics* vol. 10 691–703 (2009) DOI:10.1038/nrg2640.

260. Grandi, N. & Tramontano, E. Human endogenous retroviruses are ancient acquired elements still shaping innate immune responses. *Frontiers in Immunology* vol. 9 2039

(2018) DOI:10.3389/fimmu.2018.02039.

261. Griffiths, D. J. Endogenous retroviruses in the human genome sequence. *Genome Biology* vol. 2 reviews1017.1 (2001) DOI:10.1186/gb-2001-2-6-reviews1017.

262. Lander, E. S. *et al.* Initial sequencing and analysis of the human genome. *Nature* **409**, 860–921 (2001) DOI:10.1038/35057062.

263. Skala, A. M. Retroviral DNA Transposition: Themes and Variations. *Microbiol. Spectr.* **2**, (2014) DOI:10.1128/microbiolspec.MDNA3-0005-2014.

264. Su, Q. *et al.* A DNA transposon-based approach to validate oncogenic mutations in the mouse. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 19904–19909 (2008) DOI:10.1073/pnas.0807785105.

265. Nusse, R. The activation of cellular oncogenes by retroviral insertion. *Trends Genet.* **2**, 244–247 (1986) DOI:10.1016/0168-9525(86)90247-7.

266. Montagna, M. *et al.* Identification of a 3 kb Alu-mediated BRCA1 gene rearrangement in two breast/ovarian cancer families. *Oncogene* **18**, 4160–4165 (1999) DOI:10.1038/sj.onc.1202754.

267. Teugels, E. *et al.* De novo Alu element insertions targeted to a sequence common to the BRCA1 and BRCA2 genes. *Hum. Mutat.* **26**, 284 (2005) DOI:10.1002/humu.9366.

268. Jiang, J.-C. & Upton, K. R. Human transposons are an abundant supply of transcription factor binding sites and promoter activities in breast cancer cell lines. *Mob. DNA* **10**, 16 (2019) DOI:10.1186/s13100-019-0158-3.

269. McEachern, M. J. & Blackburn, E. H. Cap-prevented recombination between terminal telomeric repeat arrays (telomere CPR) maintains telomeres in Kluyveromyces lactis lacking telomerase. *Genes Dev.* **10**, 1822–1834 (1996) DOI:10.1101/gad.10.14.1822.

270. Maser, R. S. & DePinho, R. A. Telomeres and the DNA damage response: Why the fox is guarding the henhouse. *DNA Repair* vol. 3 979–988 (2004) DOI:10.1016/j.dnarep.2004.05.009.

271. Kupiec, M. Biology of telomeres: Lessons from budding yeast. *FEMS Microbiology Reviews* vol. 38 144–171 (2014) DOI:10.1111/1574-6976.12054.

272. Rai, R. & Chang, S. Monitoring the DNA damage response at dysfunctional telomeres. in *Methods in Molecular Biology* vol. 1343 175–180 (Humana Press Inc., 2015). DOI:10.1007/978-1-4939-2963-4_14.

273. Kurzhals, R. L. *et al.* Chromosome healing is promoted by the telomere cap component hiphop in drosophila. *Genetics* **207**, 949–959 (2017) DOI:10.1534/genetics.117.300317.

274. Westhorpe, F. G. & Straight, A. F. The centromere: Epigenetic control of chromosome segregation during mitosis. *Cold Spring Harb. Perspect. Biol.* **7**, (2015) DOI:10.1101/cshperspect.a015818.

275. Wang, L. I., Das, A. & McKim, K. S. Sister centromere fusion during meiosis I depends on maintaining cohesins and destabilizing microtubule attachments. *PLoS Genet.* **15**, (2019) DOI:10.1371/journal.pgen.1008072.

276. Sharma, G. G. *et al.* Human Heterochromatin Protein 1 Isoforms HP1 Hsα and HP1 Hsβ Interfere with hTERT-Telomere Interactions and Correlate with Changes in Cell Growth and Response to Ionizing Radiation . *Mol. Cell. Biol.* **23**, 8363–8376 (2003) DOI:10.1128/mcb.23.22.8363-8376.2003.

277. De Lange, T. Shelterin: The protein complex that shapes and safeguards human telomeres. *Genes and Development* vol. 19 2100–2110 (2005) DOI:10.1101/gad.1346005.

278. Fluminhan, A., De Aguiar-Perecin, M. L. R. & Dos Santos, J. A. Evidence for heterochromatin involvement in chromosome breakage in maize callus culture. *Ann. Bot.* **78**, 73–81 (1996) DOI:10.1006/anbo.1996.0098.

279. Okita, A. K. *et al.* Heterochromatin suppresses gross chromosomal rearrangements at centromeres by repressing Tfs1/TFIIS-dependent transcription. *Commun. Biol.* **2**, 1–13 (2019) DOI:10.1038/s42003-018-0251-z.

280. Bernard, P. *et al.* Requirement of heterochromatin for cohesion at centromeres. *Science*

(80-.). **294**, 2539–2542 (2001) DOI:10.1126/science.1064027.

281. Dernburg, A. F., Sedat, J. W. & Hawley, R. S. Direct evidence of a role for heterochromatin in meiotic chromosome segregation. *Cell* **86**, 135–46 (1996) DOI:10.1016/s0092-8674(00)80084-7.

282. Sato, K., Imai, T., Okayasu, R. & Shimokawa, T. Heterochromatin Domain Number Correlates with X-Ray and Carbon-Ion Radiation Resistance in Cancer Cells. *Radiat. Res.* **182**, 408 (2014) DOI:10.1667/rr13492.1.

283. Takata, H. *et al.* Chromatin Compaction Protects Genomic DNA from Radiation Damage. *PLoS One* **8**, (2013) DOI:10.1371/journal.pone.0075622.

284. Nakamura, S. *et al.* Bmi1 confers resistance to oxidative stress on hematopoietic stem cells. *PLoS One* **7**, e36209 (2012) DOI:10.1371/journal.pone.0036209.

285. Abdouh, M. *et al.* Bmi1 is down-regulated in the aging brain and displays antioxidant and protective activities in neurons. *PLoS One* **7**, e31870 (2012) DOI:10.1371/journal.pone.0031870.

286. Chatoo, W. *et al.* The Polycomb Group Gene Bmi1 Regulates Antioxidant Defenses in Neurons by Repressing p53 Pro-Oxidant Activity. *J. Neurosci.* **29**, 529–542 (2009) DOI:10.1523/jneurosci.5303-08.2009.

287. Westphal, T. & Reuter, G. Recombinogenic effects of suppressors of position-effect variegation in Drosophila. *Genetics* (2002).

288. Kim, J. A., Kruhlak, M., Dotiwala, F., Nussenzweig, A. & Haber, J. E. Heterochromatin is refractory to γ-H2AX modification in yeast and mammals. *J. Cell Biol.* **178**, 209–218 (2007) DOI:10.1083/jcb.200612031.

289. Cowell, I. G. *et al.* γH2AX Foci Form Preferentially in Euchromatin after Ionising-Radiation. *PLoS One* **2**, e1057 (2007) DOI:10.1371/journal.pone.0001057.

290. Peng, J. C. & Karpen, G. H. H3K9 methylation and RNA interference regulate nucleolar organization and repeated DNA stability. *Nat. Cell Biol.* **9**, 25–35 (2007)

DOI:10.1038/ncb1514.

291. Jakob, B. *et al.* DNA double-strand breaks in heterochromatin elicit fast repair protein recruitment, histone H2AX phosphorylation and relocation to euchromatin. *Nucleic Acids Res.* **39**, 6489–6499 (2011) DOI:10.1093/nar/gkr230.

292. Chiolo, I. *et al.* Double-strand breaks in heterochromatin move outside of a dynamic HP1a domain to complete recombinational repair. *Cell* **144**, 732–744 (2011) DOI:10.1016/j.cell.2011.02.012.

293. Zeller, P. *et al.* Histone H3K9 methylation is dispensable for Caenorhabditis elegans development but suppresses RNA:DNA hybrid-associated repeat instability. *Nat. Genet.* **48**, 1385–1395 (2016) DOI:10.1038/ng.3672.

294. Zeller, P. & Gasser, S. M. The Importance of Satellite Sequence Repression for Genome Stability. *Cold Spring Harb. Symp. Quant. Biol.* **82**, 15–24 (2017) DOI:10.1101/sqb.2017.82.033662.

295. Allshire, R. C. & Madhani, H. D. Ten principles of heterochromatin formation and function. *Nature Reviews Molecular Cell Biology* vol. 19 229–244 (2018) DOI:10.1038/nrm.2017.119.

296. Dixon, J. R. *et al.* Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* **485**, 376–380 (2012) DOI:10.1038/nature11082.

297. Pombo, A. & Dillon, N. Three-dimensional genome architecture: Players and mechanisms. *Nature Reviews Molecular Cell Biology* vol. 16 245–257 (2015) DOI:10.1038/nrm3965.

298. Yu, M. & Ren, B. The Three-Dimensional Organization of Mammalian Genomes. *Annu. Rev. Cell Dev. Biol.* **33**, 265–289 (2017) DOI:10.1146/annurev-cellbio-100616-060531.

299. Kalverda, B., Röling, M. D. & Fornerod, M. Chromatin organization in relation to the nuclear periphery. *FEBS Letters* vol. 582 2017–2022 (2008) DOI:10.1016/j.febslet.2008.04.015.

300. Rae, M. M. & Franke, W. W. The interphase distribution of satellite DNA-containing heterochromatin in mouse nuclei. *Chromosoma* **39**, 443–56 (1972) DOI:10.1007/bf00326177.

301. Caron, H. *et al.* The human transcriptome map: Clustering of highly expressed genes in chromosomal domains. *Science (80-.).* **291**, 1289–1293 (2001) DOI:10.1126/science.1056794.

302. Mattei, M. ; L. J. Heterochromatin, from chromosome to protein. *Atlas Genet Cytogenet Oncol Haematol. 2003;7(2)135-143.* (2003).

303. Matharu, N. & Ahituv, N. Minor Loops in Major Folds: Enhancer–Promoter Looping, Chromatin Restructuring, and Their Association with Transcriptional Regulation and Disease. *PLoS Genetics* vol. 11 (2015) DOI:10.1371/journal.pgen.1005640.

304. Navneet Matharu, N. A. File:Structural organization of chromatin.png - Wikimedia Commons. *Wikimedia Commons* https://commons.wikimedia.org/wiki/File:Structural_organization_of_chromatin.png (2016).

305. Muller, H. J. Types of visible variations induced by X-rays in Drosophila. *J. Genet.* **22**, 299–334 (1930) DOI:10.1007/BF02984195.

306. Muller HJ. Further studies on the nature and causes of gene mutations. . *Proc. Sixth Int. Congr. Genet. Ithaca, New York 1 213–255. - Open Access Libr.* (1932).

307. Weiler, K. S. & Wakimoto, B. T. Chromosome rearrangements induce both variegated and reduced, uniform expression of heterochromatic genes in a development-specific manner. *Genetics* **149**, 1451–64 (1998).

308. Wakimoto, B. T. Beyond the nucleosome: epigenetic aspects of position-effect variegation in Drosophila. *Cell* **93**, 321–4 (1998) DOI:10.1016/s0092-8674(00)81159-9.

309. Akhtar, W. *et al.* Chromatin position effects assayed by thousands of reporters integrated in parallel. *Cell* **154**, 914–927 (2013) DOI:10.1016/j.cell.2013.07.018.

310. Liu, C. Strategies for designing transgenic DNA constructs. *Methods Mol. Biol.* **1027**, 183–201 (2013) DOI:10.1007/978-1-60327-369-5_8.

311. Timms, R. T., Tchasovnikarova, I. A. & Lehner, P. J. Position-effect variegation revisited:

HUSHing up heterochromatin in human cells. *BioEssays* **38**, 333–343 (2016) DOI:10.1002/bies.201500184.

312. Trojer, P. & Reinberg, D. Facultative Heterochromatin: Is There a Distinctive Molecular Signature? *Molecular Cell* vol. 28 1–13 (2007) DOI:10.1016/j.molcel.2007.09.011.

313. Lewis, E. B. A gene complex controlling segmentation in Drosophila. *Nature* vol. 276 565–570 (1978) DOI:10.1038/276565a0.

314. Celniker, S. E., Keelan, D. J. & Lewis, E. B. The molecular genetics of the bithorax complex of Drosophila: characterization of the products of the Abdominal-B domain. *Genes Dev.* **3**, 1424–1436 (1989) DOI:10.1101/gad.3.9.1424.

315. Peifer, M., Karch, F. & Bender, W. The bithorax complex: control of segmental identity. *Genes & Dev.* **1**, 891–898 (1987) DOI:10.1101/gad.1.9.891.

316. Heuer, J. G. & Kaufman, T. C. Homeotic genes have specific functional roles in the establishment of the Drosophila embryonic peripheral nervous system. *Development* **115**, 35–47 (1992).

317. Brunk, B. P., Martin, E. C. & Adler, P. N. Drosophila genes Posterior Sex Combs and Suppressor two of zeste encode proteins with homology to the murine bmi-1 oncogene. *Nature* **353**, 351–353 (1991) DOI:10.1038/353351a0.

318. Pearce, J. J., Singh, P. B. & Gaunt, S. J. The mouse has a Polycomb-like chromobox gene. *Development* **114**, 921–9 (1992).

319. Muller, J., Gaunt, S. & Lawrence, P. A. Function of the Polycomb protein is conserved in mice and flies. *Development* **121**, 2847–2852 (1995).

320. Schuettengruber, B., Chourrout, D., Vervoort, M., Leblanc, B. & Cavalli, G. Genome Regulation by Polycomb and Trithorax Proteins. *Cell* vol. 128 735–745 (2007) DOI:10.1016/j.cell.2007.02.009.

321. Levine, S. S. *et al.* The Core of the Polycomb Repressive Complex Is Compositionally and Functionally Conserved in Flies and Humans. *Mol. Cell. Biol.* **22**, 6070–6078 (2002)

DOI:10.1128/mcb.22.17.6070-6078.2002.

322. Pasini, D., Bracken, A. P., Hansen, J. B., Capillo, M. & Helin, K. The Polycomb Group Protein Suz12 Is Required for Embryonic Stem Cell Differentiation. *Mol. Cell. Biol.* **27**, 3769–3779 (2007) DOI:10.1128/mcb.01432-06.

323. 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) DOI:10.1101/gad.381706.

324. Morey, L. *et al.* Polycomb Regulates Mesoderm Cell Fate-Specification in Embryonic Stem Cells through Activation and Repression Mechanisms. *Cell Stem Cell* **17**, 300–315 (2015) DOI:10.1016/j.stem.2015.08.009.

325. Lee, T. I. *et al.* Control of Developmental Regulators by Polycomb in Human Embryonic Stem Cells. *Cell* **125**, 301–313 (2006) DOI:10.1016/j.cell.2006.02.043.

326. Boyer, L. A. *et al.* Polycomb complexes repress developmental regulators in murine embryonic stem cells. *Nature* **441**, 349–353 (2006) DOI:10.1038/nature04733.

327. Jacobs, J. L. J. L. *et al.* The oncogene and Polycombgroup gene bmi-1 regulates cell proliferation and senescence through the ink4a locus. *Nature* (1999) doi:10.1038/16476 DOI:10.1038/16476.

328. Agherbi, H. *et al.* Polycomb mediated epigenetic silencing and replication timing at the INK4a/ARF locus during senescence. *PLoS One* **4**, (2009) DOI:10.1371/journal.pone.0005622.

329. Guo, W. J., Datta, S., Band, V. & Dimri, G. P. Mel-18, a polycomb group protein, regulates cell proliferation and senescence via transcriptional repression of Bmi-1 and c-Myc oncoproteins. *Mol. Biol. Cell* **18**, 536–546 (2007) DOI:10.1091/mbc.E06-05-0447.

330. Luis, N. M. *et al.* Regulation of human epidermal stem cell proliferation and senescence requires polycomb-dependent and independent functions of cbx4. *Cell Stem Cell* **9**, 233–246 (2011) DOI:10.1016/j.stem.2011.07.013.

331. de Napoles, M. *et al.* Polycomb group proteins ring1A/B link ubiquitylation of histone H2A to heritable gene silencing and X inactivation. *Dev. Cell* **7**, 663–676 (2004) DOI:10.1016/j.devcel.2004.10.005.

332. Brockdorff, N. Polycomb complexes in X chromosome inactivation. *Philosophical Transactions of the Royal Society B: Biological Sciences* vol. 372 (2017) DOI:10.1098/rstb.2017.0021.

333. Wang, J. *et al.* Imprinted X inactivation maintained by a mouse Polycomb group gene. *Nat. Genet.* **28**, 371–375 (2001) DOI:10.1038/ng574.

334. Bernstein, E. *et al.* Mouse Polycomb Proteins Bind Differentially to Methylated Histone H3 and RNA and Are Enriched in Facultative Heterochromatin. *Mol. Cell. Biol.* **26**, 2560–2569 (2006) DOI:10.1128/mcb.26.7.2560-2569.2006.

335. Sparmann, A. & Van Lohuizen, M. Polycomb silencers control cell fate, development and cancer. *Nature Reviews Cancer* vol. 6 846–856 (2006) DOI:10.1038/nrc1991.

336. Wang, W. *et al.* Polycomb Group (PcG) Proteins and Human Cancers: Multifaceted Functions and Therapeutic Implications. *Med. Res. Rev.* **35**, 1220–1267 (2015) DOI:10.1002/med.21358.

337. Poynter, S. T. & Kadoch, C. Polycomb and trithorax opposition in development and disease. *Wiley Interdisciplinary Reviews: Developmental Biology* vol. 5 659–688 (2016) DOI:10.1002/wdev.244.

338. Laible, G. *et al.* Mammalian homologues of the Polycomb-group gene Enhancer of zeste mediate gene silencing in Drosophila heterochromatin and at S. cerevisiae telomeres. *EMBO J.* **16**, 3219–3232 (1997) DOI:10.1093/emboj/16.11.3219.

339. 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) DOI:10.1371/journal.pone.0084147.

340. Jiao, L. & Liu, X. Structural basis of histone H3K27 trimethylation by an active polycomb

repressive complex 2. *Science (80-.).* **350**, (2015) DOI:10.1126/science.aac4383.

341. Pasini, D., Bracken, A. P., Jensen, M. R., Denchi, E. L. & Helin, K. Suz12 is essential for mouse development and for EZH2 histone methyltransferase activity. *EMBO J.* **23**, 4061–4071 (2004) DOI:10.1038/sj.emboj.7600402.

342. Nekrasov, M., Wild, B. & Müller, J. Nucleosome binding and histone methyltransferase activity of Drosophila PRC2. *EMBO Rep.* **6**, 348–353 (2005) DOI:10.1038/sj.embor.7400376.

343. Ketel, C. S. *et al.* Subunit contributions to histone methyltransferase activities of fly and worm polycomb group complexes. *Mol. Cell. Biol.* **25**, 6857–68 (2005) DOI:10.1128/MCB.25.16.6857-6868.2005.

344. 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) DOI:10.1101/gad.1035902.

345. Czermin, B. *et al.* Drosophila enhancer of Zeste/ESC complexes have a histone H3 methyltransferase activity that marks chromosomal Polycomb sites. *Cell* **111**, 185–196 (2002) DOI:10.1016/S0092-8674(02)00975-3.

346. 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) DOI:10.1016/j.cub.2005.04.051.

347. Song, J. J., Garlick, J. D. & Kingston, R. E. Structural basis of histone H4 recognition by p55. *Genes Dev.* **22**, 1313–1318 (2008) DOI:10.1101/gad.1653308.

348. Mejetta, S. *et al.* Jarid2 regulates mouse epidermal stem cell activation and differentiation. *EMBO J.* **30**, 3635–3646 (2011) DOI:10.1038/emboj.2011.265.

349. Li, G. *et al.* Jarid2 and PRC2, partners in regulating gene expression. *Genes Dev.* **24**, 368–380 (2010) DOI:10.1101/gad.1886410.

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

DOI:10.1016/j.molcel.2014.12.020.

351. Margueron, R. & Reinberg, D. The Polycomb complex PRC2 and its mark in life. *Nature* vol. 469 343–349 (2011) DOI:10.1038/nature09784.

352. 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) DOI:10.1073/pnas.1814634116.

353. Ma, R. G., Zhang, Y., Sun, T. T. & Cheng, B. Epigenetic regulation by polycomb group complexes: Focus on roles of CBX proteins. *Journal of Zhejiang University: Science B* vol. 15 412–428 (2014) DOI:10.1631/jzus.B1400077.

354. Kaustov, L. *et al.* Recognition and specificity determinants of the human Cbx chromodomains. (2010) doi:10.1074/jbc.M110.191411 DOI:10.1074/jbc.M110.191411.

355. Fang, J., Chen, T., Chadwick, B., Li, E. & Zhang, Y. Ring1b-mediated H2A ubiquitination associates with inactive X chromosomes and is involved in initiation of X inactivation. *J. Biol. Chem.* **279**, 52812–52815 (2004) DOI:10.1074/jbc.C400493200.

356. Sewalt, R. G. A. B. *et al.* Selective Interactions between Vertebrate Polycomb Homologs and the SUV39H1 Histone Lysine Methyltransferase Suggest that Histone H3-K9 Methylation Contributes to Chromosomal Targeting of Polycomb Group Proteins. *Mol. Cell. Biol.* **22**, 5539–5553 (2002) DOI:10.1128/mcb.22.15.5539-5553.2002.

357. Maison, C., Bailly, D., Quivy, J. P. & Almouzni, G. The methyltransferase Suv39h1 links the SUMO pathway to HP1α marking at pericentric heterochromatin. *Nat. Commun.* **7**, 1–9 (2016) DOI:10.1038/ncomms12224.

358. Al-Sady, B., Madhani, H. D. & Narlikar, G. J. Division of labor between the chromodomains of HP1 and Suv39 methylase enables coordination of heterochromatin spread. *Mol. Cell* **51**, 80–91 (2013) DOI:10.1016/j.molcel.2013.06.013.

359. Wang, T. *et al.* Crystal Structure of the Human SUV39H1 Chromodomain and Its Recognition of Histone H3K9me2/3. *PLoS One* **7**, 52977 (2012)

DOI:10.1371/journal.pone.0052977.

360. Firestein, R., Cui, X., Huie, P. & Cleary, M. L. Set Domain-Dependent Regulation of Transcriptional Silencing and Growth Control by SUV39H1, a Mammalian Ortholog ofDrosophila Su(var)3-9. *Mol. Cell. Biol.* **20**, 4900–4909 (2000) DOI:10.1128/mcb.20.13.4900-4909.2000.

361. Müller, M. M., Fierz, B., Bittova, L., Liszczak, G. & Muir, T. W. A two-state activation mechanism controls the histone methyltransferase Suv39h1. *Nat. Chem. Biol.* **12**, 188–193 (2016) DOI:10.1038/nchembio.2008.

362. Jehanno, C., Flouriot, G., Le Goff, P. & Michel, D. A model of dynamic stability of H3K9me3 heterochromatin to explain the resistance to reprogramming of differentiated cells *. doi:10.1016/j.bbagrm.2016.11.006 DOI:10.1016/j.bbagrm.2016.11.006.

363. Lachner, M., O'Carroll, D., Rea, S., Mechtler, K. & Jenuwein, T. Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins. *Nature* **410**, 116–120 (2001) DOI:10.1038/35065132.

364. Bannister, A. J. *et al.* Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain. *Nature* **410**, 120–124 (2001) DOI:10.1038/35065138.

365. Canzio, D. *et al.* A conformational switch in HP1 releases auto-inhibition to drive heterochromatin assembly. *Nature* **496**, 377–381 (2013) DOI:10.1038/nature12032.

366. Danzer, J. R. & Wallrath, L. L. Mechanisms of HP1-mediated gene silencing in Drosophila. *Development* **131**, 3571–3580 (2004) DOI:10.1242/dev.01223.

367. Vissers, J. H. A., Van Lohuizen, M. & Citterio, E. The emerging role of Polycomb repressors in the response to DNA damage. *Journal of Cell Science* vol. 125 3939–3948 (2012) DOI:10.1242/jcs.107375.

368. Merini, W. & Calonje, M. PRC1 is taking the lead in PcG repression. *Plant J.* **83**, 110–120 (2015) DOI:10.1111/tpj.12818.

369. Schoeftner, S. *et al.* Recruitment of PRC1 function at the initiation of X inactivation

independent of PRC2 and silencing. **25**, 3110–3122 (2006) DOI:10.1038/sj.emboj.7601187.

370. Tavares, L. *et al.* RYBP-PRC1 complexes mediate H2A ubiquitylation at polycomb target sites independently of PRC2 and H3K27me3. *Cell* **148**, 664–678 (2012) DOI:10.1016/j.cell.2011.12.029.

371. Gao, Z. *et al.* PCGF Homologs, CBX Proteins, and RYBP Define Functionally Distinct PRC1 Family Complexes. *Mol. Cell* **45**, 344–356 (2012) DOI:10.1016/j.molcel.2012.01.002.

372. Yu, M. *et al.* Direct Recruitment of Polycomb Repressive Complex 1 to Chromatin by Core Binding Transcription Factors. *Mol. Cell* **45**, 330–343 (2012) DOI:10.1016/j.molcel.2011.11.032.

373. Blackledge, N. P. *et al.* Variant PRC1 complex-dependent H2A ubiquitylation drives PRC2 recruitment and polycomb domain formation. *Cell* **157**, 1445–1459 (2014) DOI:10.1016/j.cell.2014.05.004.

374. Cooper, S. *et al.* Targeting polycomb to pericentric heterochromatin in embryonic stem cells reveals a role for H2AK119u1 in PRC2 recruitment. *Cell Rep.* **7**, 1456–70 (2014) DOI:10.1016/j.celrep.2014.04.012.

375. Kalb, R. *et al.* Histone H2A monoubiquitination promotes histone H3 methylation in Polycomb repression. *Nat. Struct. Mol. Biol.* **21**, 569–571 (2014) DOI:10.1038/nsmb.2833.

376. Cooper, S. *et al.* Jarid2 binds mono-ubiquitylated H2A lysine 119 to mediate crosstalk between Polycomb complexes PRC1 and PRC2. *Nat. Commun.* **7**, 1–8 (2016) DOI:10.1038/ncomms13661.

377. Tamburri, S. *et al.* Histone H2AK119 Mono-Ubiquitination Is Essential for Polycomb-Mediated Transcriptional Repression. *Mol. Cell* **77**, 840-856.e5 (2020) DOI:10.1016/j.molcel.2019.11.021.

378. Fursova, N. A. *et al.* Synergy between Variant PRC1 Complexes Defines Polycomb-Mediated Gene Repression. *Mol. Cell* **74**, 1020-1036.e8 (2019) DOI:10.1016/j.molcel.2019.03.024.

379. Moussa, H. F. *et al.* Canonical PRC1 controls sequence-independent propagation of

Polycomb-mediated gene silencing. *Nat. Commun.* **10**, 1–12 (2019) DOI:10.1038/s41467-019-09628-6.

380. Blackledge, N. P. *et al.* PRC1 Catalytic Activity Is Central to Polycomb System Function. *Mol. Cell* **77**, 857-874.e9 (2020) DOI:10.1016/j.molcel.2019.12.001.

381. Vidal, M. Role of polycomb proteins Ring1A and Ring1Bin the epigenetic regulation of gene expression. *Int. J. Dev. Biol.* **53**, 355–370 (2009) DOI:10.1387/ijdb.082690mv.

382. van Lohuizen, M. *et al.* Identification of cooperating oncogenes in Eμ-myc transgenic mice by provirus tagging. *Cell* **65**, 737–752 (1991) DOI:10.1016/0092-8674(91)90382-9.

383. Haupt, Y., Alexander, W. S., Barri, G., Klinken, S. P. & Adams, J. M. Novel zinc finger gene implicated as myc collaborator by retrovirally accelerated lymphomagenesis in E mu-myc transgenic mice. *Cell* **65**, 753–63 (1991) DOI:10.1016/0092-8674(91)90383-a.

384. Bhattacharya, R., Banerjee Mustafi, S., Street, M., Dey, A. & Dwivedi, S. K. D. Bmi-1: At the crossroads of physiological and pathological biology. *Genes and Diseases* vol. 2 225–239 (2015) DOI:10.1016/j.gendis.2015.04.001.

385. Cohen, K. J., Hanna, J. S., Prescott, J. E. & Dang, C. V. Transformation by the Bmi-1 oncoprotein correlates with its subnuclear localization but not its transcriptional suppression activity. *Mol. Cell. Biol.* **16**, 5527–5535 (1996) DOI:10.1128/mcb.16.10.5527.

386. Yadav, A. K. *et al.* Deletion analysis of BMI1 oncoprotein identifies its negative regulatory domain. *Mol. Cancer* **9**, 158 (2010) DOI:10.1186/1476-4598-9-158.

387. Zhang, J. & Sarge, K. D. Identification of a polymorphism in the RING finger of human Bmi-1 that causes its degradation by the ubiquitin-proteasome system. *FEBS Lett.* **583**, 960–964 (2009) DOI:10.1016/j.febslet.2009.02.023.

388. Balasubramanian, S., Scharadin, T. M., Han, B., Xu, W. & Eckert, R. L. The Bmi-1 helix-turn and ring finger domains are required for bmi-1 antagonism of (-) epigallocatechin-3-gallate suppression of skin cancer cell survival. *Cell. Signal.* **27**, 1336–1344 (2015) DOI:10.1016/j.cellsig.2015.03.021.

389. Satijn, D. P. E. & Otte, A. P. Polycomb group protein complexes: Do different complexes regulate distinct target genes? *Biochimica et Biophysica Acta - Gene Structure and Expression* vol. 1447 1–16 (1999) DOI:10.1016/S0167-4781(99)00130-X.

390. Satijn, D. P. E. & Otte, A. P. RING1 Interacts with Multiple Polycomb-Group Proteins and Displays Tumorigenic Activity. *Mol. Cell. Biol.* **19**, 57–68 (1999) DOI:10.1128/mcb.19.1.57.

391. Li, Z. *et al.* Structure of a Bmi-1-Ring1B polycomb group ubiquitin ligase complex. *J. Biol. Chem.* **281**, 20643–20649 (2006) DOI:10.1074/jbc.M602461200.

392. Buchwald, G. *et al.* Structure and E3-ligase activity of the Ring-Ring complex of Polycomb proteins Bmi1 and Ring1b. *EMBO J.* **25**, 2465–2474 (2006) DOI:10.1038/sj.emboj.7601144.

393. Bentley, M. L. *et al.* Recognition of UbcH5c and the nucleosome by the Bmi1/Ring1b ubiquitin ligase complex. *EMBO J.* **30**, 3285–3297 (2011) DOI:10.1038/emboj.2011.243.

394. Gray, F. *et al.* BMI1 regulates PRC1 architecture and activity through homo- and hetero-oligomerization. *Nat. Commun.* **7**, 13343 (2016) DOI:10.1038/ncomms13343.

395. Chagraoui, J. *et al.* E4F1: A novel candidate factor for mediating BMI1 function in primitive hematopoietic cells. *Genes Dev.* **20**, 2110–2120 (2006) DOI:10.1101/gad.1453406.

396. NCBI Gene. BMI1 BMI1 proto-oncogene, polycomb ring finger [Homo sapiens (human)] - Gene - NCBI. https://www.ncbi.nlm.nih.gov/gene/648.

397. UniProtKB/Swiss-Prot., U. P. R. (UniProt). Polycomb complex protein BMI-1 - Homo sapiens (Human). https://www.uniprot.org/uniprot/P35226.

398. the human Protein Atlas. Tissue expression of BMI1 - Summary - The Human Protein Atlas. https://www.proteinatlas.org/ENSG00000168283-BMI1/tissue.

399. Tang, Z. *et al.* GEPIA. http://gepia.cancer-pku.cn/detail.php?gene=bmi1.

400. Tang, Z. *et al.* GEPIA: A web server for cancer and normal gene expression profiling and interactive analyses. *Nucleic Acids Res.* **45**, W98–W102 (2017) DOI:10.1093/nar/gkx247.

401. Zhu, S. *et al.* BMI1 regulates androgen receptor in prostate cancer independently of the

polycomb repressive complex 1. *Nat. Commun.* **9**, (2018) DOI:10.1038/s41467-018-02863-3.

402. Zhu, S. *et al.* BMI1 is directly regulated by androgen receptor to promote castration-resistance in prostate cancer. *Oncogene* **39**, 17–29 (2020) DOI:10.1038/s41388-019-0966-4.

403. Hu, J. *et al.* BTF3 sustains cancer stem-like phenotype of prostate cancer via stabilization of BMI1. *J. Exp. Clin. Cancer Res.* **38**, 227 (2019) DOI:10.1186/s13046-019-1222-z.

404. Yin, T. *et al.* Bmi1 inhibition enhances the sensitivity of pancreatic cancer cells to gemcitabine. *Oncotarget* **7**, 37192–37204 (2016) DOI:10.18632/oncotarget.9293.

405. Karimi Mazraehshah, M. *et al.* Anticancer effects of miR-200c in colorectal cancer through BMI1. *J. Cell. Biochem.* **119**, 10005–10012 (2018) DOI:10.1002/jcb.27330.

406. Tsai, H.-L. *et al.* Impact of BMI1 expression on the apoptotic effect of paclitaxel in colorectal cancer. *Am. J. Cancer Res.* **9**, 2544–2553 (2019).

407. Yu, F., Zhou, C., Zeng, H., Liu, Y. & Li, S. BMI1 activates WNT signaling in colon cancer by negatively regulating the WNT antagonist IDAX. *Biochem. Biophys. Res. Commun.* **496**, 468–474 (2018) DOI:10.1016/j.bbrc.2018.01.063.

408. Crunkhorn, S. Cancer: BMI1 inhibition reverses lung cancer. *Nat. Rev. Drug Discov.* **15**, 678 (2016) DOI:10.1038/nrd.2016.197.

409. Vonlanthen, S. *et al.* The bmi-1 oncoprotein is differentially expressed in non-small cell lung cancer and correlates with INK4A-ARF locus expression. *Br. J. Cancer* **84**, 1372–6 (2001) DOI:10.1054/bjoc.2001.1791.

410. Li, W. *et al.* Repression of Noxa by Bmi1 contributes to deguelin-induced apoptosis in non-small cell lung cancer cells. *J. Cell. Mol. Med.* **22**, 6213–6227 (2018) DOI:10.1111/jcmm.13908.

411. Liu, F. *et al.* microRNA 141 represses nasopharyngeal carcinoma growth through inhibiting BMI1. *Oncol. Lett.* **16**, 6479–6487 (2018) DOI:10.3892/ol.2018.9444.

412. Wu, J. *et al.* Down-regulation of BMI-1 cooperates with artemisinin on growth inhibition of nasopharyngeal carcinoma cells. *J. Cell. Biochem.* **112**, 1938–1948 (2011) DOI:10.1002/jcb.23114.

413. Hu, Q. *et al.* The poor outcome of second primary oral squamous cell carcinoma is attributed to Bmi1 upregulation. *Cancer Med.* **7**, 1056–1069 (2018) DOI:10.1002/cam4.1348.

414. Srinivasan, M. *et al.* Downregulation of Bmi1 in breast cancer stem cells suppresses tumor growth and proliferation. *Oncotarget* **8**, 38731–38742 (2017) DOI:10.18632/oncotarget.16317.

415. M., J. R. & Vaishnave, S. BMI1 and PTEN are key determinants of breast cancer therapy: A plausible therapeutic target in breast cancer. *Gene* vol. 678 302–311 (2018) DOI:10.1016/j.gene.2018.08.022.

416. Behesti, H., Bhagat, H., Dubuc, A. M., Taylor, M. D. & Marino, S. Bmi1 overexpression in the cerebellar granule cell lineage of mice affects cell proliferation and survival without initiating medulloblastoma formation. *DMM Dis. Model. Mech.* **6**, 49–63 (2013) DOI:10.1242/dmm.009506.

417. Bakhshinyan, D. *et al.* BMI1 is a therapeutic target in recurrent medulloblastoma. *Oncogene* **38**, 1702–1716 (2019) DOI:10.1038/s41388-018-0549-9.

418. Badodi, S. *et al.* Convergence of BMI1 and CHD7 on ERK Signaling in Medulloblastoma. *Cell Rep.* **21**, 2772–2784 (2017) DOI:10.1016/j.celrep.2017.11.021.

419. Merve, A. *et al.* Polycomb group gene bMI1 controls invasion of medulloblastoma cells and inhibits BMP-regulated cell adhesion. *Acta Neuropathol. Commun.* **2**, 10 (2014) DOI:10.1186/2051-5960-2-10.

420. Nowak, K. *et al.* BMI1 is a target gene of E2F-1 and is strongly expressed in primary neuroblastomas. *Nucleic Acids Res.* **34**, 1745–1754 (2006) DOI:10.1093/nar/gkl119.

421. Cui, H. *et al.* Bmi-1 is essential for the tumorigenicity of neuroblastoma cells. *Am. J. Pathol.*

170, 1370–1378 (2007) DOI:10.2353/ajpath.2007.060754.

422. Cui, H. *et al.* Bmi-1 regulates the differentiation and clonogenic self-renewal of I-type neuroblastoma cells in a concentration-dependent manner. *J. Biol. Chem.* **281**, 34696–704 (2006) DOI:10.1074/jbc.M604009200.

423. Facchino, S., Abdouh, M., Chatoo, W. & Bernier, G. BMI1 confers radioresistance to normal and cancerous neural stem cells through recruitment of the DNA damage response machinery. *J. Neurosci.* **30**, 10096–111 (2010) DOI:10.1523/JNEUROSCI.1634-10.2010.

424. Hong, Y., Shang, C., Xue, Y. & Liu, Y. Silencing of Bmi-1 gene enhances chemotherapy sensitivity in human glioblastoma cells. *Med. Sci. Monit.* **21**, 1002–7 (2015) DOI:10.12659/MSM.893754.

425. Abdouh, M. *et al.* BMI1 sustains human glioblastoma multiforme stem cell renewal. *J. Neurosci.* **29**, 8884–96 (2009) DOI:10.1523/JNEUROSCI.0968-09.2009.

426. Flamier, A. *et al.* Off-target effect of the BMI1 inhibitor PTC596 drives epithelial-mesenchymal transition in glioblastoma multiforme. *npj Precis. Oncol.* **4**, 1 (2020) DOI:10.1038/s41698-019-0106-1.

427. Itahana, K. *et al.* Control of the Replicative Life Span of Human Fibroblasts by p16 and the Polycomb Protein Bmi-1. *Mol. Cell. Biol.* **23**, 389–401 (2003) DOI:10.1128/mcb.23.1.389-401.2003.

428. Dimri, G. P. *et al.* The Bmi-1 oncogene induces telomerase activity and immortalizes human mammary epithelial cells. *Cancer Res.* **62**, 4736–45 (2002).

429. Liu, L., Andrews, L. G. & Tollefsbol, T. O. Loss of the human polycomb group protein BMI1 promotes cancer-specific cell death. *Oncogene* **25**, 4370–4375 (2006) DOI:10.1038/sj.onc.1209454.

430. Jacobs, J. L., Kieboom, K., Marino, S., DePinho, R. A. & Van Lohuizen, M. The oncogene and Polycombgroup gene bmi-1 regulates cell proliferation and senescence through the ink4a locus. *Nature* **397**, 164–168 (1999) DOI:10.1038/16476.

431. McConnell, B. B., Gregory, F. J., Stott, F. J., Hara, E. & Peters, G. Induced Expression of p16 INK4a Inhibits Both CDK4- and CDK2-Associated Kinase Activity by Reassortment of Cyclin-CDK-Inhibitor Complexes . *Mol. Cell. Biol.* **19**, 1981–1989 (1999) DOI:10.1128/mcb.19.3.1981.

432. Topacio, B. R. *et al.* Cyclin D-Cdk4,6 Drives Cell-Cycle Progression via the Retinoblastoma Protein's C-Terminal Helix. *Mol. Cell* **74**, 758-770.e4 (2019) DOI:10.1016/j.molcel.2019.03.020.

433. Lees, J. A. *et al.* The retinoblastoma protein binds to a family of E2F transcription factors. *Mol. Cell. Biol.* **13**, 7813–7825 (1993) DOI:10.1128/mcb.13.12.7813.

434. Fuchs, S. Y., Adler, V., Buschmann, T., Wu, X. & Ev Ronai, Z. '. *Mdm2 association with p53 targets its ubiquitination*. http://www.stockton-press.co.uk/onc.

435. Brooks, C. L. & Gu, W. p53 ubiquitination: Mdm2 and beyond. *Molecular Cell* vol. 21 307–315 (2006) DOI:10.1016/j.molcel.2006.01.020.

436. Calao, M. *et al.* Direct effects of Bmi1 on p53 protein stability inactivates oncoprotein stress responses in embryonal cancer precursor cells at tumor initiation. *Oncogene* **32**, 3616–3626 (2013) DOI:10.1038/onc.2012.368.

437. Shen, J. *et al.* The E3 Ligase RING1 Targets p53 for Degradation and Promotes Cancer Cell Proliferation and Survival. (2017) doi:10.1158/0008-5472.CAN-17-1805 DOI:10.1158/0008-5472.CAN-17-1805.

438. Spring, L., Bardia, A. & Modi, S. Targeting the cyclin d-cyclin-dependent kinase (CDK) 4/6-retinoblastoma pathway with selective cdk 4/6 inhibitors in hormone receptor-positive breast cancer: Rationale, current status, and future directions. *Discov. Med.* **21**, (2016).

439. Wade, M. & Wahl, G. M. Targeting Mdm2 and mdmx in cancer therapy: Better living through medicinal chemistry? *Molecular Cancer Research* vol. 7 1–11 (2009) DOI:10.1158/1541-7786.MCR-08-0423.

440. Alajez, N. M. *et al.* Targeted depletion of BMI1 sensitizes tumor cells to P53-mediated

apoptosis in response to radiation therapy. *Cell Death Differ.* **16**, 1469–1479 (2009) DOI:10.1038/cdd.2009.85.

441. Leung, C. *et al.* Bmi1 is essential for cerebellar development and is overexpressed in human medulloblastomas. *Nature* (2004) doi:10.1038/nature02385 DOI:10.1038/nature02385.

442. Van der Lugt, N. M. T. *et al.* Posterior transformation, neurological abnormalities, and severe hematopoietic defects in mice with a targeted deletion of the bmi-1 proto-oncogene. *Genes Dev.* **8**, 757–769 (1994) DOI:10.1101/gad.8.7.757.

443. Barabino, A. *et al.* Loss of Bmi1 causes anomalies in retinal development and degeneration of cone photoreceptors. *Development* **143**, 1571–1584 (2016) DOI:10.1242/dev.125351.

444. Dhawan, S., Tschen, S. I. & Bhushan, A. Bmi-1 regulates the Ink4a/Arf locus to control pancreatic β-cell proliferation. *Genes Dev.* **23**, 906–911 (2009) DOI:10.1101/gad.1742609.

445. Gu, M. *et al.* Heterozygous knockout of the Bmi-1 gene causes an early onset of phenotypes associated with brain aging. *Age (Omaha).* **36**, 129–139 (2014) DOI:10.1007/s11357-013-9552-9.

446. Cordisco, S. *et al.* Bmi-1 reduction plays a key role in physiological and premature aging of primary human keratinocytes. *J. Invest. Dermatol.* **130**, 1048–1062 (2010) DOI:10.1038/jid.2009.355.

447. El Hajjar, J. *et al.* Heterochromatic genome instability and neurodegeneration sharing similarities with Alzheimer's disease in old Bmi1+/− mice. *Sci. Rep.* **9**, 594 (2019) DOI:10.1038/s41598-018-37444-3.

448. Flamier, A. *et al.* Modeling Late-Onset Sporadic Alzheimer's Disease through BMI1 Deficiency. *Cell Rep.* **23**, 2653–2666 (2018) DOI:10.1016/j.celrep.2018.04.097.

449. Nishibuchi, G. & Déjardin, J. The molecular basis of the organization of repetitive DNA-containing constitutive heterochromatin in mammals. *Chromosome Research* vol. 25 77–87 (2017) DOI:10.1007/s10577-016-9547-3.

450. Ugarković, D. & Plohl, M. Variation in satellite DNA profiles - Causes and effects. *EMBO*

Journal vol. 21 5955–5959 (2002) DOI:10.1093/emboj/cdf612.

451. Black, E. M. & Giunta, S. Repetitive fragile sites: Centromere satellite DNA as a source of genome instability in human diseases. *Genes* vol. 9 (2018) DOI:10.3390/genes9120615.

452. Madireddy, A. & Gerhardt, J. Replication through repetitive DNA elements and their role in human diseases. in *Advances in Experimental Medicine and Biology* vol. 1042 549–581 (Springer New York LLC, 2017). DOI:10.1007/978-981-10-6955-0_23.

453. Gadgil, R., Barthelemy, J., Lewis, T. & Leffak, M. Replication stalling and DNA microsatellite instability. *Biophys. Chem.* **225**, 38–48 (2017) DOI:10.1016/j.bpc.2016.11.007.

454. Warmerdam, D. O. & Wolthuis, R. M. F. Keeping ribosomal DNA intact: a repeating challenge. *Chromosome Research* vol. 27 57–72 (2019) DOI:10.1007/s10577-018-9594-z.

455. Dunn, M. J. & Anderson, M. Z. To repeat or not to repeat: Repetitive sequences regulate genome stability in candida albicans. *Genes* vol. 10 (2019) DOI:10.3390/genes10110866.

456. Bettin, N., Oss Pegorar, C. & Cusanelli, E. The Emerging Roles of TERRA in Telomere Maintenance and Genome Stability. *Cells* **8**, 246 (2019) DOI:10.3390/cells8030246.

457. Yunis, J. J. & Yasmineh, W. G. Heterochromatin, satellite DNA, and cell function. Structural DNA of eucaryotes may support and protect genes and aid in speciation. *Science* **174**, 1200–1209 (1971).

458. Louzada, S. *et al.* Decoding the role of satellite DNA in genome architecture and plasticity—an evolutionary and clinical affair. *Genes* vol. 11 (2020) DOI:10.3390/genes11010072.

459. Lee, H., Zhang, Z. & Krause, H. M. Long Noncoding RNAs and Repetitive Elements: Junk or Intimate Evolutionary Partners? *Trends in Genetics* vol. 35 892–902 (2019) DOI:10.1016/j.tig.2019.09.006.

460. Bouchard, R. A. Moderately repetitive DNA in evolution. *Int. Rev. Cytol.* **76**, 113–93 (1982) DOI:10.1016/s0074-7696(08)61790-8.

461. Smit, A. RepeatMasker. *Nature* https://media.nature.com/full/nature-

assets/ni/journal/v5/n8/extref/repeatmasker.html.

462. Charlesworth, B., Sniegowski, P. & Stephan, W. The evolutionary dynamics of repetitive DNA in eukaryotes. *Nature* vol. 371 215–220 (1994) DOI:10.1038/371215a0.

463. Gregory, R. T. *The Evolution of the Genome*. *The Evolution of the Genome* (Elsevier Inc., 2005). doi:10.1016/B978-0-12-301463-4.X5000-1 DOI:10.1016/B978-0-12-301463-4.X5000-1.

464. Levinson, G. & Gutman, G. Slipped-strand mispairing: a major mechanism for DNA sequence evolution. *Mol. Biol. Evol.* (1987) doi:10.1093/oxfordjournals.molbev.a040442 DOI:10.1093/oxfordjournals.molbev.a040442.

465. Wilder, J. & Hollocher, H. Mobile elements and the genesis of microsatellites in Dipterans. *Mol. Biol. Evol.* **18**, 384–392 (2001) DOI:10.1093/oxfordjournals.molbev.a003814.

466. Plohl, M., Meštrović, N. & Mravinac, B. Satellite DNA evolution. *Genome Dyn.* **7**, 126–152 (2012) DOI:10.1159/000337122.

467. Plohl, M., Luchetti, A., Meštrović, N. & Mantovani, B. Satellite DNAs between selfishness and functionality: Structure, genomics and evolution of tandem repeats in centromeric (hetero)chromatin. *Gene* **409**, 72–82 (2008) DOI:10.1016/j.gene.2007.11.013.

468. Plohl, M., Meštrović, N. & Mravinac, B. Centromere identity from the DNA point of view. *Chromosoma* vol. 123 313–325 (2014) DOI:10.1007/s00412-014-0462-0.

469. Biscotti, M. A., Canapa, A., Forconi, M., Olmo, E. & Barucca, M. Transcription of tandemly repetitive DNA: functional roles. *Chromosom. Res.* **23**, 463–477 (2015) DOI:10.1007/s10577-015-9494-4.

470. Ravindran, S. Barbara McClintock and the discovery of jumping genes. *Proc. Natl. Acad. Sci. U. S. A.* **109**, 20198–20199 (2012) DOI:10.1073/pnas.1219372109.

471. Hickman, A. B. & Dyda, F. DNA Transposition at Work. *Chemical Reviews* vol. 116 12758–12784 (2016) DOI:10.1021/acs.chemrev.6b00003.

472. Mita, P. & Boeke, J. D. How retrotransposons shape genome regulation. *Current Opinion in Genetics and Development* vol. 37 90–100 (2016) DOI:10.1016/j.gde.2016.01.001.

473. Wagstaff, B. J., Barnerßoi, M. & Roy-Engel, A. M. Evolutionary conservation of the functional modularity of primate and murine LINE-1 elements. *PLoS One* **6**, (2011) DOI:10.1371/journal.pone.0019672.

474. Martin, S. L. The ORF1 Protein Encoded by LINE-1: Structure and Function During L1 Retrotransposition. *J. Biomed. Biotechnol.* **2006**, (2006) DOI:10.1155/JBB/2006/45621.

475. Holmes, S. E., Singer, M. F. & Swergold, G. D. Studies on p40, the leucine zipper motif-containing protein encoded by the first open reading frame of an active human LINE-1 transposable element. *J. Biol. Chem.* **267**, 19765–8 (1992).

476. Naufer, M. N., Furano, A. V. & Williams, M. C. Protein-nucleic acid interactions of LINE-1 ORF1p. *Seminars in Cell and Developmental Biology* vol. 86 140–149 (2019) DOI:10.1016/j.semcdb.2018.03.019.

477. Feng, Q., Moran, J. V., Kazazian, H. H. & Boeke, J. D. Human L1 retrotransposon encodes a conserved endonuclease required for retrotransposition. *Cell* **87**, 905–916 (1996) DOI:10.1016/S0092-8674(00)81997-2.

478. Mathias, S. L., Scott, A. F., Kazazian, H. H., Boeke, J. D. & Gabriel, A. Reverse transcriptase encoded by a human transposable element. *Science (80-.).* **254**, 1808–1810 (1991) DOI:10.1126/science.1722352.

479. Kines, K. J. *et al.* The endonuclease domain of the LINE-1 ORF2 protein can tolerate multiple mutations. *Mob. DNA* **7**, 8 (2016) DOI:10.1186/s13100-016-0064-x.

480. Cost, G. J. & Boeke, J. D. Targeting of human retrotransposon integration is directed by the specificity of the L1 endonuclease for regions of unusual DNA structure. *Biochemistry* **37**, 18081–18093 (1998) DOI:10.1021/bi981858s.

481. Wimmer, K., Callens, T., Wernstedt, A. & Messiaen, L. The NF1 gene contains hotspots for L1 endonuclease-dependent De Novo insertion. *PLoS Genet.* **7**, (2011)

DOI:10.1371/journal.pgen.1002371.

482. Muotri, A. R. *et al.* Somatic mosaicism in neuronal precursor cells mediated by L1 retrotransposition. *Nature* **435**, 903–910 (2005) DOI:10.1038/nature03663.

483. Singer, T., McConnell, M. J., Marchetto, M. C. N., Coufal, N. G. & Gage, F. H. LINE-1 retrotransposons: Mediators of somatic variation in neuronal genomes? *Trends Neurosci.* **33**, 345–354 (2010) DOI:10.1016/j.tins.2010.04.001.

484. Wallace, N., Wagstaff, B. J., Deininger, P. L. & Roy-Engel, A. M. LINE-1 ORF1 protein enhances Alu SINE retrotransposition. *Gene* **419**, 1–6 (2008) DOI:10.1016/j.gene.2008.04.007.

485. Grechishnikova, D. & Poptsova, M. Conserved 3' UTR stem-loop structure in L1 and Alu transposons in human genome: Possible role in retrotransposition. *BMC Genomics* **17**, 992 (2016) DOI:10.1186/s12864-016-3344-4.

486. Vidaud, D. *et al.* Haemophilia B due to a de novo insertion of a human-specific Alu subfamily member within the coding region of the factor IX gene. *Eur. J. Hum. Genet.* **1**, 30–36 (1993) DOI:10.1159/000472385.

487. Sukarova, E., Dimovski, A. J., Tchacarova, P., Petkov, G. H. & Efremov, G. D. An Alu Insert as the Cause of a Severe Form of Hemophilia A. *Acta Haematol.* **106**, 126–129 (2001) DOI:10.1159/000046602.

488. Ganguly, A., Dunbar, T., Chen, P., Godmilow, L. & Ganguly, T. Exon skipping caused by an intronic insertion of a young Alu Yb9 element leads to severe hemophilia A. *Hum. Genet.* **113**, 348–352 (2003) DOI:10.1007/s00439-003-0986-5.

489. Chen, J. M. *et al.* Detection of two Alu insertions in the CFTR gene. *J. Cyst. Fibros.* **7**, 37–43 (2008) DOI:10.1016/j.jcf.2007.04.001.

490. Campbell, S. Will Biotechnology Stop Aging? : the quest to slow, stop, or turn back the body's clock. *IEEE Pulse* **10**, 3–7 (2019) DOI:10.1109/MPULS.2019.2899701.

491. Eissenberg, J. C. Hungering for Immortality. *Mo. Med.* **115**, 12–17 (2018).

492. Blagosklonny, M. V. Rapamycin for longevity: Opinion article. *Aging (Albany. NY).* **11**, 8048–8067 (2019) DOI:10.18632/aging.102355.

493. Gurău, F. *et al.* Anti-senescence compounds: A potential nutraceutical approach to healthy aging. *Ageing Research Reviews* vol. 46 14–31 (2018) DOI:10.1016/j.arr.2018.05.001.

494. Carrel, A. On the permanent life of tissues outside of the organism. *J. Exp. Med.* **15**, 516–528 (1912) DOI:10.1084/jem.15.5.516.

495. Hayflick, L. The establishment of a line (WISH) of human amnion cells in continuous cultivation. *Exp. Cell Res.* **23**, 14–20 (1961) DOI:10.1016/0014-4827(61)90059-3.

496. Hayflick, L. & Moorhead, P. S. The serial cultivation of human diploid cell strains. *Exp. Cell Res.* **25**, 585–621 (1961) DOI:10.1016/0014-4827(61)90192-6.

497. Burnet, M. & Burnet, M. Intrinsic Mutagenesis: An Elementary Discussion at the Level of Molecular Biology. in *Intrinsic mutagenesis* 10–16 (Springer Netherlands, 1974). doi:10.1007/978-94-011-6606-5_2 DOI:10.1007/978-94-011-6606-5_2.

498. Masoro, E. J. & Austad, S. N. *Handbook of the Biology of Aging*. *Handbook of the Biology of Aging* (Elsevier Inc., 2011). doi:10.1016/C2009-0-01967-9 DOI:10.1016/C2009-0-01967-9.

499. Coppé, J. P. *et al.* Senescence-associated secretory phenotypes reveal cell-nonautonomous functions of oncogenic RAS and the p53 tumor suppressor. *PLoS Biol.* **6**, (2008) DOI:10.1371/journal.pbio.0060301.

500. Campisi, J. *et al.* From discoveries in ageing research to therapeutics for healthy ageing. *Nature* vol. 571 183–192 (2019) DOI:10.1038/s41586-019-1365-2.

501. Wright, W. E. & Hayflick, L. The regulation of cellular aging by nuclear events in cultured normal human fibroblasts (WI-38). *Adv. Exp. Med. Biol.* **61**, 39–55 (1975) DOI:10.1007/978-1-4615-9032-3_3.

502. Wright, W. E. & Hayflick, L. Nuclear control of cellular aging demonstrated by hybridization of anucleate and whole cultured normal human fibroblasts. *Exp. Cell Res.* **96**, 113–121 (1975) DOI:10.1016/S0014-4827(75)80043-7.

503. Allsopp, R. C. *et al.* Telomere Shortening Is Associated with Cell Division in Vitro and in Vivo. *Exp. Cell Res.* **220**, 194–200 (1995) DOI:10.1006/excr.1995.1306.

504. Hastie, N. D. *et al.* Telomere reduction in human colorectal carcinoma and with ageing. *Nature* **346**, 866–868 (1990) DOI:10.1038/346866a0.

505. Harley, C. B. Telomere loss: mitotic clock or genetic time bomb? *Mutat. Res. DNAging* **256**, 271–282 (1991) DOI:10.1016/0921-8734(91)90018-7.

506. Chang, E. & Harley, C. B. Telomere length and replicative aging in human vascular tissues. *Proc. Natl. Acad. Sci. U. S. A.* **92**, 11190–11194 (1995) DOI:10.1073/pnas.92.24.11190.

507. Van Steensel, B. & De Lange, T. Control of telomere length by the human telomeric protein TRF1. *Nature* **385**, 740–743 (1997) DOI:10.1038/385740a0.

508. Allsopp, R. C. *et al.* Telomere length predicts replicative capacity of human fibroblasts. *Proc. Natl. Acad. Sci. U. S. A.* **89**, 10114–10118 (1992) DOI:10.1073/pnas.89.21.10114.

509. Bodnar, A. G. *et al.* Extension of life-span by introduction of telomerase into normal human cells. *Science (80-.).* **279**, 349–352 (1998) DOI:10.1126/science.279.5349.349.

510. Lansdorp, P. M. *et al.* Heterogeneity in telomere length of human chromosomes. *Hum. Mol. Genet.* **5**, 685–91 (1996) DOI:10.1093/hmg/5.5.685.

511. McElligott, R. & Wellinger, R. J. The terminal DNA structure of mammalian chromosomes. *EMBO J.* **16**, 3705–14 (1997) DOI:10.1093/emboj/16.12.3705.

512. Griffith, J. D. *et al.* Mammalian telomeres end in a large duplex loop. *Cell* **97**, 503–514 (1999) DOI:10.1016/S0092-8674(00)80760-6.

513. Xin, H., Liu, D. & Songyang, Z. The telosome/shelterin complex and its functions. *Genome biology* vol. 9 232 (2008) DOI:10.1186/gb-2008-9-9-232.

514. Van Steensel, B., Smogorzewska, A. & De Lange, T. TRF2 protects human telomeres from end-to-end fusions. *Cell* **92**, 401–413 (1998) DOI:10.1016/S0092-8674(00)80932-0.

515. Smogorzewska, A., Karlseder, J., Holtgreve-Grez, H., Jauch, A. & De Lange, T. DNA ligase IV-

dependent NHEJ of deprotected mammalian telomeres in G1 and G2. *Curr. Biol.* **12**, 1635–1644 (2002) DOI:10.1016/S0960-9822(02)01179-X.

516. Karlseder, J. *et al.* The telomeric protein TRF2 binds the ATM Kinase and Can Inhibit the ATM-dependent DNA damage response. *PLoS Biol.* **2**, (2004) DOI:10.1371/journal.pbio.0020240.

517. Denchi, E. L. & De Lange, T. Protection of telomeres through independent control of ATM and ATR by TRF2 and POT1. *Nature* **448**, 1068–1071 (2007) DOI:10.1038/nature06065.

518. Wills, E. J. The powerhouse of the cell. *Ultrastructural Pathology* vol. 16 iii–vi (1992) DOI:10.3109/01913129209061353.

519. Sulek, K. [Nobel prize to Hans Adolf Krebs for discovery of the citric acid cycle and to Fritz Albert Lipmann in 1953 for discovery of coenzyme A and its importance in intermediary metabolism]. *Wiad. Lek.* **21**, 2187–9 (1968).

520. Turrens, J. F. Mitochondrial formation of reactive oxygen species. *Journal of Physiology* vol. 552 335–344 (2003) DOI:10.1113/jphysiol.2003.049478.

521. Cheeseman, K. H. & Slater, T. F. An introduction to free radical biochemistry. *Br. Med. Bull.* **49**, 481–493 (1993) DOI:10.1093/oxfordjournals.bmb.a072625.

522. Holley, A. K., Bakthavatchalu, V., Velez-Roman, J. M. & St. Clair, D. K. Manganese superoxide dismutase: Guardian of the powerhouse. *International Journal of Molecular Sciences* vol. 12 7114–7162 (2011) DOI:10.3390/ijms12107114.

523. Lubos, E., Loscalzo, J. & Handy, D. E. Glutathione peroxidase-1 in health and disease: From molecular mechanisms to therapeutic opportunities. *Antioxidants and Redox Signaling* vol. 15 1957–1997 (2011) DOI:10.1089/ars.2010.3586.

524. Couto, N., Wood, J. & Barber, J. The role of glutathione reductase and related enzymes on cellular redox homoeostasis network. *Free Radical Biology and Medicine* vol. 95 27–42 (2016) DOI:10.1016/j.freeradbiomed.2016.02.028.

525. Stanley, B. A. *et al.* Thioredoxin reductase-2 is essential for keeping low levels of H 2O 2

emission from isolated heart mitochondria. *J. Biol. Chem.* **286**, 33669–33677 (2011) DOI:10.1074/jbc.M111.284612.

526. Lu, J. & Holmgren, A. The thioredoxin antioxidant system. *Free Radical Biology and Medicine* vol. 66 75–87 (2014) DOI:10.1016/j.freeradbiomed.2013.07.036.

527. Glorieux, C. & Calderon, P. B. Catalase, a remarkable enzyme: Targeting the oldest antioxidant enzyme to find a new cancer treatment approach. *Biological Chemistry* vol. 398 1095–1108 (2017) DOI:10.1515/hsz-2017-0131.

528. Apostolova, N. & Victor, V. M. Molecular strategies for targeting antioxidants to mitochondria: Therapeutic implications. *Antioxidants and Redox Signaling* vol. 22 686–729 (2015) DOI:10.1089/ars.2014.5952.

529. Horan, M. P., Pichaud, N. & Ballard, J. W. O. Review: Quantifying mitochondrial dysfunction in complex diseases of aging. *Journals Gerontol. - Ser. A Biol. Sci. Med. Sci.* **67 A**, 1022–1035 (2012) DOI:10.1093/gerona/glr263.

530. Pichaud, N. *et al.* Age Dependent Dysfunction of Mitochondrial and ROS Metabolism Induced by Mitonuclear Mismatch. *Front. Genet.* **10**, 130 (2019) DOI:10.3389/fgene.2019.00130.

531. Yun, J. & Finkel, T. Mitohormesis. *Cell Metabolism* vol. 19 757–766 (2014) DOI:10.1016/j.cmet.2014.01.011.

532. Schriner, S. E. *et al.* Medecine: Extension of murine life span by overexpression of catalase targeted to mitochondria. *Science (80-.).* **308**, 1909–1911 (2005) DOI:10.1126/science.1106653.

533. Treuting, P. M. *et al.* Reduction of age-associated pathology in old mice by overexpression of catalase in mitochondria. *J. Gerontol. A. Biol. Sci. Med. Sci.* **63**, 813–22 (2008) DOI:10.1093/gerona/63.8.813.

534. Dai, D. F. *et al.* Overexpression of catalase targeted to mitochondria attenuates murine cardiac aging. *Circulation* **119**, 2789–2797 (2009)

DOI:10.1161/CIRCULATIONAHA.108.822403.

535. Guo, C. Y., Sun, L., Chen, X. P. & Zhang, D. S. Oxidative stress, mitochondrial damage and neurodegenerative diseases. *Neural Regen. Res.* **8**, 2003–2014 (2013) DOI:10.3969/j.issn.1673-5374.2013.21.009.

536. Lin, M. T. & Beal, M. F. Mitochondrial dysfunction and oxidative stress in neurodegenerative diseases. *Nature* vol. 443 787–795 (2006) DOI:10.1038/nature05292.

537. Andersen, J. K. Oxidative stress in neurodegeneration: Cause or consequence? *Nat. Rev. Neurosci.* **10**, S18 (2004) DOI:10.1038/nrn1434.

538. Stefanatos, R. & Sanz, A. The role of mitochondrial ROS in the aging brain. *FEBS Letters* vol. 592 743–758 (2018) DOI:10.1002/1873-3468.12902.

539. Reddy, P. H. Mitochondrial oxidative damage in aging and Alzheimer's disease: implications for mitochondrially targeted antioxidant therapeutics. *J. Biomed. Biotechnol.* **2006**, 31372 (2006) DOI:10.1155/JBB/2006/31372.

540. Zindy, F., Quelle, D. E., Roussel, M. F. & Sherr, C. J. Expression of the p16(INK4a) tumor suppressor versus other INK4 family members during mouse development and aging. *Oncogene* **15**, 203–211 (1997) DOI:10.1038/sj.onc.1201178.

541. Kotake, Y., Naemura, M., Murasaki, C., Inoue, Y. & Okamoto, H. Transcriptional regulation of the p16 tumor suppressor gene. *Anticancer Research* vol. 35 4397–4402 (2015).

542. Oberdoerffer, P. & Sinclair, D. A. The role of nuclear architecture in genomic instability and ageing. *Nature Reviews Molecular Cell Biology* vol. 8 692–702 (2007) DOI:10.1038/nrm2238.

543. Haithcock, E. *et al.* Age-related changes of nuclear architecture in Caenorhabditis elegans. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 16690–16695 (2005) DOI:10.1073/pnas.0506955102.

544. Wood, J. G. *et al.* Chromatin remodeling in the aging genome of Drosophila. *Aging Cell* **9**, 971–978 (2010) DOI:10.1111/j.1474-9726.2010.00624.x.

545. Frankel, S., Woods, J., Ziafazeli, T. & Rogina, B. RPD3 histone deacetylase and nutrition have distinct but interacting effects on Drosophila longevity. *Aging (Albany. NY).* **7**, 1112–1129 (2015) DOI:10.18632/aging.100856.

546. Li, W. *et al.* Activation of transposable elements during aging and neuronal decline in Drosophila. *Nat. Neurosci.* **16**, 529–531 (2013) DOI:10.1038/nn.3368.

547. Gaubatz, J. W. & Cutler, R. G. Mouse satellite DNA is transcribed in senescent cardiac muscle. *J. Biol. Chem.* **265**, 17753–8 (1990).

548. Scaffidi, P. & Misteli, T. Lamin A-dependent nuclear defects in human aging. *Science (80-.).* **312**, 1059–63 (2006) DOI:10.1126/science.1127168.

549. Fraga, M. F. *et al.* Epigenetic differences arise during the lifetime of monozygotic twins. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 10604–10609 (2005) DOI:10.1073/pnas.0500398102.

550. Herskind, A. M. *et al.* The heritability of human longevity: A population-based study of 2872 Danish twin pairs born 1870-1900. *Hum. Genet.* **97**, 319–323 (1996) DOI:10.1007/BF02185763.

551. Poulsen, P., Esteller, M., Vaag, A. & Fraga, M. F. The epigenetic basis of twin discordance in age-related diseases. *Pediatr. Res.* **61**, (2007) DOI:10.1203/pdr.0b013e31803c7b98.

552. Pal, S. & Tyler, J. K. Epigenetics and aging. *Science Advances* vol. 2 (2016) DOI:10.1126/sciadv.1600584.

553. Csoka, A. B. *et al.* Novel lamin A/C gene (LMNA) mutations in atypical progeroid syndromes. *J. Med. Genet.* **41**, 304–308 (2004) DOI:10.1136/jmg.2003.015651.

554. Cao, H. & Hegele, R. A. LMNA is mutated in Hutchinson-Gilford progeria (MIM 176670) but not in Wiedemann-Rautenstrauch progeroid syndrome (MIM 264090). *J. Hum. Genet.* **48**, 271–274 (2003) DOI:10.1007/s10038-003-0025-3.

555. Hennekam, R. C. M. Hutchinson-Gilford progeria syndrome: Review of the phenotype. in *American Journal of Medical Genetics, Part A* vol. 140 2603–2624 (Am J Med Genet A, 2006). DOI:10.1002/ajmg.a.31346.

556. Wallace, R. G. *et al.* The role of epigenetics in cardiovascular health and ageing: A focus on physical activity and nutrition. *Mechanisms of Ageing and Development* vol. 174 76–85 (2018) DOI:10.1016/j.mad.2017.11.013.

557. Kochmanski, J. *et al.* Longitudinal effects of developmental bisphenol A and variable diet exposures on epigenetic drift in mice. *Reprod. Toxicol.* **68**, 154–163 (2017) DOI:10.1016/j.reprotox.2016.07.021.

558. Adams, P. D. *et al.* Lysosome-mediated processing of chromatin in senescence. *J. Cell Biol.* **202**, 129–143 (2013) DOI:10.1083/jcb.201212110.

559. O'Sullivan, R. J., Kubicek, S., Schreiber, S. L. & Karlseder, J. Reduced histone biosynthesis and chromatin changes arising from a damage signal at telomeres. *Nat. Struct. Mol. Biol.* **17**, 1218–1225 (2010) DOI:10.1038/nsmb.1897.

560. Feser, J. *et al.* Elevated Histone Expression Promotes Life Span Extension. *Mol. Cell* **39**, 724–735 (2010) DOI:10.1016/j.molcel.2010.08.015.

561. Rogakou, E. P. & Sekeri-Pataryas, K. E. Histone variants of H2A and H3 families are regulated during in vitro aging in the same manner as during differentiation. *Exp. Gerontol.* **34**, 741–754 (1999) DOI:10.1016/S0531-5565(99)00046-7.

562. Zee, B. M., Levin, R. S., Dimaggio, P. A. & Garcia, B. A. Global turnover of histone post-translational modifications and variants in human cells. *Epigenetics and Chromatin* **3**, (2010) DOI:10.1186/1756-8935-3-22.

563. Das, C. & Tyler, J. K. Histone exchange and histone modifications during transcription and aging. *Biochim. Biophys. Acta - Gene Regul. Mech.* **1819**, 332–342 (2012) DOI:10.1016/j.bbagrm.2011.08.001.

564. Tvardovskiy, A., Schwämmle, V., Kempf, S. J., Rogowska-Wrzesinska, A. & Jensen, O. N. Accumulation of histone variant H3.3 with age is associated with profound changes in the histone methylation landscape. *Nucleic Acids Res.* (2017) doi:10.1093/nar/gkx696 DOI:10.1093/nar/gkx696.

565. Lindahl, T. Instability and decay of the primary structure of DNA. *Nature* vol. 362 709–715 (1993) DOI:10.1038/362709a0.

566. Martin, L. J. DNA Damage and Repair. *J. Neuropathol. Exp. Neurol.* **67**, 377–387 (2008) DOI:10.1097/NEN.0b013e31816ff780.

567. Soares, J. P. *et al.* Aging and DNA damage in humans: A meta-analysis study. *Aging (Albany. NY).* **6**, 432–439 (2014) DOI:10.18632/aging.100667.

568. Wu, Y. & Brosh Jr., R. FANCJ Helicase Operates in the Fanconi Anemia DNA Repair Pathway and the Response to Replicational Stress. *Curr. Mol. Med.* **9**, 470–482 (2009) DOI:10.2174/156652409788167159.

569. Suhasini, A. N. & Brosh, R. M. Fanconi anemia and Bloom's syndrome crosstalk through FANCJ-BLM helicase interaction. *Trends in Genetics* vol. 28 7–13 (2012) DOI:10.1016/j.tig.2011.09.003.

570. Black, J. O. Xeroderma Pigmentosum. *Head Neck Pathol.* **10**, 139–144 (2016) DOI:10.1007/s12105-016-0707-8.

571. Sugitani, N., Sivley, R. M., Perry, K. E., Capra, J. A. & Chazin, W. J. XPA: A key scaffold for human nucleotide excision repair. *DNA Repair* vol. 44 123–135 (2016) DOI:10.1016/j.dnarep.2016.05.018.

572. Nair, N., Shoaib, M. & Sørensen, C. S. Chromatin Dynamics in Genome Stability: Roles in Suppressing Endogenous DNA Damage and Facilitating DNA Repair. *Int. J. Mol. Sci.* **18**, 1486 (2017) DOI:10.3390/ijms18071486.

573. World Health Organisation. *Global action plan on the public health response to dementia 2017 - 2025*. *Geneva: World Health Organization* http://www.who.int/mental_health/neurology/dementia/action_plan_2017_2025/en/ (2017).

574. Hippius, H. & Neundörfer, G. The discovery of Alzheimer's disease. *Dialogues Clin. Neurosci.* **5**, 101–108 (2003).

575. Alzheimer, A. Uber einen eigenartigen schweren Erkrankungsprozess der Hirninde. *Neurol. Cent.* (1906).

576. Ziegler-Graham, K., Brookmeyer, R., Johnson, E. & Arrighi, H. M. Worldwide variation in the doubling time of Alzheimer's disease incidence rates. *Alzheimer's Dement.* **4**, 316–323 (2008) DOI:10.1016/j.jalz.2008.05.2479.

577. Association, A. 2018 Alzheimer's disease facts and figures. *Alzheimer's Dement.* **14**, 367–429 (2018) DOI:10.1016/j.jalz.2018.02.001.

578. International, D. *World Alzheimer Report 2018 - The state of the art of dementia research: New frontiers; World Alzheimer Report 2018 - The state of the art of dementia research: New frontiers*. (2018).

579. Vespa, J., Armstrong, D. M. & Medina, L. *Demographic Turning Points for the United States: Population Projections for 2020 to 2060 Population Estimates and Projections Current Population Reports*. www.census.gov/programs-surveys/popproj (2018).

580. our world in data. Prevalence of Alzheimer disease and other dementias, World. https://ourworldindata.org/grapher/prevalence-of-dementias (2020).

581. Citron, M. *et al.* Mutation of the β-amyloid precursor protein in familial Alzheimer's disease increases β-protein production. *Nature* **360**, 672–674 (1992) DOI:10.1038/360672a0.

582. Wragg, M., Hutton, M., Talbot, C. & Goate, A. Genetic association between intronic polymorphism in presenilin-1 gene and late-onset Alzheimer's disease. *Lancet* **347**, 509–512 (1996) DOI:10.1016/S0140-6736(96)91140-X.

583. Kehoe, P., Williams, J., Lovestone, S., Wilcock, G. & Owen, M. J. Presenilin-1 polymorphism and Alzheimer's disease. The UK Alzheimer's Disease Collaborative Group. *Lancet (London, England)* **347**, 1185 (1996).

584. Deng, G., Pike, C. J. & Cotman, C. W. Alzheimer-associated presenilin-2 confers increased sensitivity to apoptosis in PC12 cells. *FEBS Lett.* **397**, 50–54 (1996) DOI:10.1016/S0014-5793(96)01142-8.

585. Scheuner, D. *et al.* Secreted amyloid β-protein similar to that in the senile plaques of Alzheimer's disease is increased in vivo by the presenilin 1 and 2 and APP mutations linked to familial Alzheimer's disease. *Nat. Med.* **2**, 864–870 (1996) DOI:10.1038/nm0896-864.

586. Keret, O., Shochat, T., Steiner, I. & Glik, A. Non-ashkenazi jewish origin is associated with early onset Alzheimer's disease. *J. Alzheimer's Dis.* **65**, 877–884 (2018) DOI:10.3233/JAD-180331.

587. van der Flier, W. M., Pijnenburg, Y. A. L., Fox, N. C. & Scheltens, P. Early-onset versus late-onset Alzheimer's disease: The case of the missing APOE ε4 allele. *The Lancet Neurology* vol. 10 280–288 (2011) DOI:10.1016/S1474-4422(10)70306-9.

588. Mendez, M. F. Early-Onset Alzheimer Disease. *Neurologic Clinics* vol. 35 263–281 (2017) DOI:10.1016/j.ncl.2017.01.005.

589. Wang, C. *et al.* Gain of toxic apolipoprotein E4 effects in human iPSC-derived neurons is ameliorated by a small-molecule structure corrector article. *Nat. Med.* **24**, 647–657 (2018) DOI:10.1038/s41591-018-0004-z.

590. Montagne, A. *et al.* APOE4 leads to blood–brain barrier dysfunction predicting cognitive decline. *Nature* 1–6 (2020) doi:10.1038/s41586-020-2247-3 DOI:10.1038/s41586-020-2247-3.

591. Iturria-Medina, Y. *et al.* Early role of vascular dysregulation on late-onset Alzheimer's disease based on multifactorial data-driven analysis. *Nat. Commun.* **7**, 1–14 (2016) DOI:10.1038/ncomms11934.

592. Sweeney, M. D., Sagare, A. P. & Zlokovic, B. V. Blood-brain barrier breakdown in Alzheimer disease and other neurodegenerative disorders. *Nature Reviews Neurology* vol. 14 133–150 (2018) DOI:10.1038/nrneurol.2017.188.

593. Nation, D. A. *et al.* Blood–brain barrier breakdown is an early biomarker of human cognitive dysfunction. *Nat. Med.* **25**, 270–276 (2019) DOI:10.1038/s41591-018-0297-y.

594. van de Haar, H. J. *et al.* Neurovascular unit impairment in early Alzheimer's disease

measured with magnetic resonance imaging. *Neurobiol. Aging* **45**, 190–196 (2016) DOI:10.1016/j.neurobiolaging.2016.06.006.

595. Jakobsdottir, J. *et al.* Rare Functional Variant in TM2D3 is Associated with Late-Onset Alzheimer's Disease. *PLoS Genet.* **12**, e1006327 (2016) DOI:10.1371/journal.pgen.1006327.

596. Mez, J. *et al.* Two novel loci, COBL and SLC10A2, for Alzheimer's disease in African Americans. *Alzheimer's Dement.* **13**, 119–129 (2017) DOI:10.1016/j.jalz.2016.09.002.

597. Korvatska, O. *et al.* R47H variant of TREM2 associated with Alzheimer disease in a large late-onset family clinical, genetic, and neuropathological study. *JAMA Neurol.* **72**, 920–927 (2015) DOI:10.1001/jamaneurol.2015.0979.

598. Conejero-Goldberg, C. *et al.* APOE2 enhances neuroprotection against alzheimer's disease through multiple molecular mechanisms. *Mol. Psychiatry* **19**, 1243–1250 (2014) DOI:10.1038/mp.2013.194.

599. Kero, M. *et al.* Amyloid precursor protein (APP) A673T mutation in the elderly Finnish population. *Neurobiol. Aging* **34**, 1518.e1-1518.e3 (2013) DOI:10.1016/j.neurobiolaging.2012.09.017.

600. Kokawa, A. *et al.* The A673T mutation in the amyloid precursor protein reduces the production of β-amyloid protein from its β-carboxyl terminal fragment in cells. *Acta Neuropathol. Commun.* **3**, 66 (2015) DOI:10.1186/s40478-015-0247-6.

601. Freudenberg-Hua, Y., Li, W. & Davies, P. The role of genetics in advancing precision medicine for Alzheimer's Disease-A narrative review. *Frontiers in Medicine* vol. 5 (2018) DOI:10.3389/fmed.2018.00108.

602. Fox, N. C., Freeborough, P. A. & Rossor, M. N. Visualisation and quantification of rates of atrophy in Alzheimer's disease. *Lancet* **348**, 94–97 (1996) DOI:10.1016/S0140-6736(96)05228-2.

603. Pini, L. *et al.* Brain atrophy in Alzheimer's Disease and aging. *Ageing Research Reviews* vol. 30 25–48 (2016) DOI:10.1016/j.arr.2016.01.002.

604. Terry, R. D. Cell Death or Synaptic Loss in Alzheimer Disease. *J. Neuropathol. Exp. Neurol.* **59**, 1118–1119 (2000) DOI:10.1093/jnen/59.12.1118.

605. McEvoy, L. K. *et al.* Alzheimer disease: Quantitative structural neuroimaging for detection and prediction of clinical and structural changes in mild cognitive impairment. *Radiology* **251**, 195–205 (2009) DOI:10.1148/radiol.2511080924.

606. Andrade-Moraes, C. H. *et al.* Cell number changes in Alzheimer's disease relate to dementia, not to plaques and tangles. *Brain* (2013) doi:10.1093/brain/awt273 DOI:10.1093/brain/awt273.

607. NIH. NIH Image Gallery | Flickr. *Flickr* https://www.flickr.com/photos/nihgov/.

608. Montenegro, J. M. F. Alzheimer's Disease Diagnosis Based on Cognitive Methods in Virtual Environments and Emotions Analysis. (2018).

609. Van Der Kant, R. & Goldstein, L. S. B. Developmental Cell Review Cellular Functions of the Amyloid Precursor Protein from Development to Dementia. *Dev. Cell* **32**, 502–515 (2015) DOI:10.1016/j.devcel.2015.01.022.

610. Kamenetz, F. *et al.* APP Processing and Synaptic Function. *Neuron* **37**, 925–937 (2003) DOI:10.1016/S0896-6273(03)00124-7.

611. Opsomer, R. *et al.* Amyloid Precursor Protein (APP) controls excitatory/inhibitory synaptic inputs by regulating the transcriptional activator Neuronal PAS Domain Protein 4 (NPAS4). *bioRxiv* 504340 (2018) doi:10.1101/504340 DOI:10.1101/504340.

612. Priller, C. *et al.* Synapse formation and function is modulated by the amyloid precursor protein. *J. Neurosci.* **26**, 7212–7221 (2006) DOI:10.1523/JNEUROSCI.1450-06.2006.

613. Plummer, S., Van Den Heuvel, C., Thornton, E., Corrigan, F. & Cappai, R. The neuroprotective properties of the amyloid precursor protein following traumatic brain injury. *Aging and Disease* vol. 7 163–179 (2016) DOI:10.14336/AD.2015.0907.

614. Tackenberg, C. & Nitsch, R. M. The secreted APP ectodomain sAPPα, but not sAPPβ, protects neurons against Aβ oligomer-induced dendritic spine loss and increased tau

phosphorylation. *Mol. Brain* **12**, 27 (2019) DOI:10.1186/s13041-019-0447-2.

615. Dar, N. J. & Glazner, G. W. Deciphering the neuroprotective and neurogenic potential of soluble amyloid precursor protein alpha (sAPPα). *Cellular and Molecular Life Sciences* 1–16 (2020) doi:10.1007/s00018-019-03404-x DOI:10.1007/s00018-019-03404-x.

616. Li, M. *et al.* The amyloid precursor protein (APP) intracellular domain regulates translation of p44, a short isoform of p53, through an IRES-dependent mechanism. *Neurobiol. Aging* **36**, 2725–2736 (2015) DOI:10.1016/j.neurobiolaging.2015.06.021.

617. Shu, R. *et al.* APP intracellular domain acts as a transcriptional regulator of miR-663 suppressing neuronal differentiation. *Cell Death Dis.* **6**, e1651–e1651 (2015) DOI:10.1038/cddis.2015.10.

618. Stevens, F. J. *et al.* A Molecular Model for Self-Assembly of Amyloid Fibrils: Immunoglobulin Light Chains. *Biochemistry* **34**, 10697–10702 (1995) DOI:10.1021/bi00034a001.

619. Cotman, C. W. The beta-Amyloid Peptide, Peptide Self-Assembly, and the Emergence of Biological Activities. *Ann. N. Y. Acad. Sci.* **814**, 1–16 (1997) DOI:10.1111/j.1749-6632.1997.tb46140.x.

620. Haass, C. & Selkoe, D. J. Soluble protein oligomers in neurodegeneration: Lessons from the Alzheimer's amyloid β-peptide. *Nature Reviews Molecular Cell Biology* vol. 8 101–112 (2007) DOI:10.1038/nrm2101.

621. Tanzi, R. E. & Bertram, L. Twenty years of the Alzheimer's disease amyloid hypothesis: A genetic perspective. *Cell* vol. 120 545–555 (2005) DOI:10.1016/j.cell.2005.02.008.

622. Mullard, A. Anti-amyloid failures stack up as Alzheimer antibody flops. *Nat. Rev. Drug Discov.* (2019) doi:10.1038/d41573-019-00064-1 DOI:10.1038/d41573-019-00064-1.

623. Cappai, R. & Barnham, K. J. Delineating the mechanism of Alzheimer's disease Aβ peptide neurotoxicity. *Neurochemical Research* vol. 33 526–532 (2008) DOI:10.1007/s11064-007-9469-8.

624. Hardy, J. The amyloid hypothesis for Alzheimer's disease: A critical reappraisal. *Journal of*

Neurochemistry vol. 110 1129–1134 (2009) DOI:10.1111/j.1471-4159.2009.06181.x.

625. Kayed, R. & Lasagna-Reeves, C. A. Molecular mechanisms of amyloid oligomers toxicity. *Journal of Alzheimer's Disease* vol. 33 S67–S78 (2013) DOI:10.3233/JAD-2012-129001.

626. Lee, H. G. *et al.* Challenging the amyloid cascade hypothesis: Senile plaques and amyloid-β as protective adaptations to Alzheimer disease. in *Annals of the New York Academy of Sciences* vol. 1019 1–4 (New York Academy of Sciences, 2004). DOI:10.1196/annals.1297.001.

627. Treusch, S., Cyr, D. M. & Lindquist, S. Amyloid deposits: Protection against toxic protein species? *Cell Cycle* vol. 8 1668–1674 (2009) DOI:10.4161/cc.8.11.8503.

628. Zhao, J., Deng, Y., Jiang, Z. & Qing, H. G protein-coupled receptors (GPCRs) in Alzheimer's disease: A focus on BACE1 related GPCRs. *Frontiers in Aging Neuroscience* vol. 8 (2016) DOI:10.3389/fnagi.2016.00058.

629. Trojanowski, J. Q., Schuck, T., Schmidt, M. L. & Lee, V. M. Distribution of tau proteins in the normal human central and peripheral nervous system. *J. Histochem. Cytochem.* **37**, 209–215 (1989) DOI:10.1177/37.2.2492045.

630. Mandelkow, E. M., Thies, E. & Mandelkow, E. Tau and axonal transport. in *Alzheimer's Disease: Advances in Genetics, Molecular and Cellular Biology* 237–256 (Springer US, 2007). doi:10.1007/978-0-387-35135-3_14 DOI:10.1007/978-0-387-35135-3_14.

631. Lee, G., Neve, R. L. & Kosik, K. S. The microtubule binding domain of tau protein. *Neuron* **2**, 1615–1624 (1989) DOI:10.1016/0896-6273(89)90050-0.

632. Kadavath, H. *et al.* Tau stabilizes microtubules by binding at the interface between tubulin heterodimers. *Proc. Natl. Acad. Sci. U. S. A.* **112**, 7501–7506 (2015) DOI:10.1073/pnas.1504081112.

633. Yu, Y. *et al.* Developmental regulation of tau phosphorylation, tau kinases, and tau phosphatases. *J. Neurochem.* **108**, 1480–1494 (2009) DOI:10.1111/j.1471-4159.2009.05882.x.

634. Lovestone, S., Hartley, C. L., Pearce, J. & Anderton, B. H. Phosphorylation of tau by glycogen synthase kinase-3β in intact mammalian cells: The effects on the organization and stability of microtubules. *Neuroscience* **73**, 1145–1157 (1996) DOI:10.1016/0306-4522(96)00126-1.

635. Gong, C. X. *et al.* Phosphorylation of microtubule-associated protein tau is regulated by protein phosphatase 2A in mammalian brain. Implications for neurofibrillary degeneration in Alzheimer's disease. *J. Biol. Chem.* **275**, 5535–5544 (2000) DOI:10.1074/jbc.275.8.5535.

636. Schönheit, B., Zarski, R. & Ohm, T. G. Spatial and temporal relationships between plaques and tangles in Alzheimer-pathology. *Neurobiology of Aging* vol. 25 697–711 (2004) DOI:10.1016/j.neurobiolaging.2003.09.009.

637. Braak, H. & Braak, E. Frequency of stages of Alzheimer-related lesions in different age categories. *Neurobiol. Aging* **18**, 351–357 (1997) DOI:10.1016/S0197-4580(97)00056-0.

638. Bretteville, A. & Planel, E. Tau aggregates: Toxic, inert, or protective species? *Journal of Alzheimer's Disease* vol. 14 431–436 (2008) DOI:10.3233/JAD-2008-14411.

639. Serrano-Pozo, A., Frosch, M. P., Masliah, E. & Hyman, B. T. Neuropathological alterations in Alzheimer disease. *Cold Spring Harb. Perspect. Med.* **1**, (2011) DOI:10.1101/cshperspect.a006189.

640. Wilcock, G. K. & Esiri, M. M. Plaques, tangles and dementia. A quantitative study. *J. Neurol. Sci.* **56**, 343–356 (1982) DOI:10.1016/0022-510X(82)90155-1.

641. Santacruz, K. *et al.* Medicine: Tau suppression in a neurodegenerative mouse model improves memory function. *Science (80-.).* **309**, 476–481 (2005) DOI:10.1126/science.1113694.

642. Kuchibhotla, K. V. *et al.* Neurofibrillary tangle-bearing neurons are functionally integrated in cortical circuits in vivo. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 510–514 (2014) DOI:10.1073/pnas.1318807111.

643. Kopeikina, K. J., Hyman, B. J. & Spires-Jones, T. L. Soluble forms of tau are toxic in alzheimer's disease. *Transl. Neurosci.* **3**, 223–233 (2012) DOI:10.2478/s13380-012-0032-y.

644. Ghag, G. *et al.* Soluble tau aggregates, not large fibrils, are the toxic species that display seeding and cross-seeding behavior. *Protein Sci.* **27**, 1901–1909 (2018) DOI:10.1002/pro.3499.

645. Lee, V. M.-Y., Goedert, M. & Trojanowski, J. Q. Neurodegenerative Tauopathies. *Annu. Rev. Neurosci.* **24**, 1121–1159 (2001) DOI:10.1146/annurev.neuro.24.1.1121.

646. Bandyopadhyay, B., Li, G., Yin, H. & Kuret, J. Tau aggregation and toxicity in a cell culture model of tauopathy. *J. Biol. Chem.* **282**, 16454–16464 (2007) DOI:10.1074/jbc.M700192200.

647. Andorfer, C. *et al.* Cell-cycle reentry and cell death in transgenic mice expressing nonmutant human tau isoforms. *J. Neurosci.* **25**, 5446–5454 (2005) DOI:10.1523/JNEUROSCI.4637-04.2005.

648. Rapoport, M., Dawson, H. N., Binder, L. I., Vitek, M. P. & Ferreira, A. Tau is essential to β-amyloid-induced neurotoxicity. *Proc. Natl. Acad. Sci. U. S. A.* **99**, 6364–6369 (2002) DOI:10.1073/pnas.092136199.

649. Roberson, E. D. *et al.* Reducing endogenous tau ameliorates amyloid β-induced deficits in an Alzheimer's disease mouse model. *Science (80-.).* **316**, 750–754 (2007) DOI:10.1126/science.1141736.

650. Lewis, J. *et al.* Enhanced neurofibrillary degeneration in transgenic mice expressing mutant tau and APP. *Science (80-.).* **293**, 1487–1491 (2001) DOI:10.1126/science.1058189.

651. Bollati, V. *et al.* DNA methylation in repetitive elements and Alzheimer disease. *Brain. Behav. Immun.* **25**, 1078–1083 (2011) DOI:10.1016/j.bbi.2011.01.017.

652. Larsen, P. A. *et al.* The Alu neurodegeneration hypothesis: A primate-specific mechanism for neuronal transcription noise, mitochondrial dysfunction, and manifestation of neurodegenerative disease. *Alzheimer's and Dementia* vol. 13 828–838 (2017) DOI:10.1016/j.jalz.2017.01.017.

653. Weissman, L. *et al.* Defective DNA base excision repair in brain from individuals with

Alzheimer's disease and amnestic mild cognitive impairment. *Nucleic Acids Res.* (2007) doi:10.1093/nar/gkm605 DOI:10.1093/nar/gkm605.

654. Yang, J.-L., Weissman, L., Bohr, V. A. & Mattson, M. P. Mitochondrial DNA damage and repair in neurodegenerative disorders. *DNA Repair (Amst).* **7**, 1110–1120 (2008) DOI:10.1016/j.dnarep.2008.03.012.

655. Hou, Y., Song, H., Croteau, D. L., Akbari, M. & Bohr, V. A. Genome instability in Alzheimer disease. *Mech. Ageing Dev.* **161**, 83–94 (2017) DOI:10.1016/j.mad.2016.04.005.

656. Shanbhag, N. M. *et al.* Early neuronal accumulation of DNA double strand breaks in Alzheimer's disease. *Acta Neuropathol. Commun.* **7**, 1–18 (2019) DOI:10.1186/s40478-019-0723-5.

657. Sultan, A. *et al.* Nuclear Tau, a key player in neuronal DNA protection. *J. Biol. Chem.* **286**, 4566–4575 (2011) DOI:10.1074/jbc.M110.199976.

658. Bou Samra, E. *et al.* A role for Tau protein in maintaining ribosomal DNA stability and cytidine deaminase-deficient cell survival. *Nat. Commun.* **8**, 1–14 (2017) DOI:10.1038/s41467-017-00633-1.

659. Iijima-Ando, K., Zhao, L. J., Gatt, A., Shenton, C. & Iijima, K. A DNA damage-activated checkpoint kinase phosphorylates tau and enhances tau-induced neurodegeneration. *Hum. Mol. Genet.* (2010) doi:10.1093/hmg/ddq068 DOI:10.1093/hmg/ddq068.

660. Frost, B., Hemberg, M., Lewis, J. & Feany, M. B. Tau promotes neurodegeneration through global chromatin relaxation. *Nat. Neurosci.* **17**, 357–366 (2014) DOI:10.1038/nn.3639.

661. Guo, C. *et al.* Tau Activates Transposable Elements in Alzheimer's Disease. *Cell Rep.* **23**, 2874–2880 (2018) DOI:10.1016/j.celrep.2018.05.004.

662. Santiard-Baron, D. *et al.* Identification of β-amyloid-responsive genes by RNA differential display: Early induction of a DNA damage-inducible gene, gadd45. *Exp. Neurol.* **158**, 206–213 (1999) DOI:10.1006/exnr.1999.7076.

663. Mao, P. & Reddy, P. H. Aging and amyloid beta-induced oxidative DNA damage and

mitochondrial dysfunction in Alzheimer's disease: Implications for early intervention and therapeutics. *Biochimica et Biophysica Acta - Molecular Basis of Disease* vol. 1812 1359–1370 (2011) DOI:10.1016/j.bbadis.2011.08.005.

664. Fodor, B. D., Shukeir, N., Reuter, G. & Jenuwein, T. Mammalian Su (var) Genes in Chromatin Control . *Annu. Rev. Cell Dev. Biol.* **26**, 471–501 (2010) DOI:10.1146/annurev.cellbio.042308.113225.

665. Kim, M., Trinh, B. N., Long, T. I., Oghamian, S. & Laird, P. W. Dnmt1 deficiency leads to enhanced microsatellite instability in mouse embryonic stem cells. *Nucleic Acids Res.* (2004) doi:10.1093/nar/gkh912 DOI:10.1093/nar/gkh912.

666. Kondo, Y. *et al.* Downregulation of Histone H3 Lysine 9 Methyltransferase G9a Induces Centrosome Disruption and Chromosome Instability in Cancer Cells. *PLoS One* **3**, e2037 (2008) DOI:10.1371/journal.pone.0002037.

667. Peng, J. C. & Karpen, G. H. Epigenetic regulation of heterochromatic DNA stability. *Current Opinion in Genetics and Development* vol. 18 204–211 (2008) DOI:10.1016/j.gde.2008.01.021.

668. Matsui, T. *et al.* Proviral silencing in embryonic stem cells requires the histone methyltransferase ESET. *Nature* **464**, 927–931 (2010) DOI:10.1038/nature08858.

669. Rowe, H. M. *et al.* KAP1 controls endogenous retroviruses in embryonic stem cells. *Nature* **463**, 237–240 (2010) DOI:10.1038/nature08674.

670. Pinheiro, I. *et al.* Prdm3 and Prdm16 are H3K9me1 methyltransferases required for mammalian heterochromatin integrity. *Cell* **150**, 948–960 (2012) DOI:10.1016/j.cell.2012.06.048.

671. Towbin, B. D. *et al.* Step-wise methylation of histone H3K9 positions heterochromatin at the nuclear periphery. *Cell* **150**, 934–947 (2012) DOI:10.1016/j.cell.2012.06.051.

672. Black, J. C. & Whetstine, J. R. Chromatin landscape: Methylation beyond transcription. *Epigenetics* **6**, 9–15 (2011) DOI:10.4161/epi.6.1.13331.

673. Wang, R. H. *et al.* Impaired DNA Damage Response, Genome Instability, and Tumorigenesis in SIRT1 Mutant Mice. *Cancer Cell* **14**, 312–323 (2008) DOI:10.1016/j.ccr.2008.09.001.

674. Kappes, F. *et al.* The DEK oncoprotein is a Su(var) that is essential to heterochromatin integrity. *Genes Dev.* **25**, 673–678 (2011) DOI:10.1101/gad.2036411.

675. Maison, C. *et al.* SUMOylation promotes de novo targeting of HP1 ± to pericentric heterochromatin. *Nat. Genet.* **43**, 220–227 (2011) DOI:10.1038/ng.765.

676. Fan, Y. *et al.* Histone H1 depletion in mammals alters global chromatin structure but causes specific changes in gene regulation. *Cell* **123**, 1199–1212 (2005) DOI:10.1016/j.cell.2005.10.028.

677. Iwase, S. *et al.* ATRX ADD domain links an atypical histone methylation recognition mechanism to human mental-retardation syndrome. *Nat. Struct. Mol. Biol.* **18**, 769–776 (2011) DOI:10.1038/nsmb.2062.

678. Maison, C. & Almouzni, G. HP1 and the dynamics of heterochromatin maintenance. *Nature Reviews Molecular Cell Biology* vol. 5 296–304 (2004) DOI:10.1038/nrm1355.

679. Peters, A. H. F. M. *et al.* Loss of the Suv39h histone methyltransferases impairs mammalian heterochromatin and genome stability. *Cell* **107**, 323–37 (2001) DOI:10.1016/S0092-8674(01)00542-6.

680. Prasanth, S. G., Shen, Z., Prasanth, K. V. & Stillman, B. Human origin recognition complex is essential for HP1 binding to chromatin and heterochromatin organization. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 15093–8 (2010) DOI:10.1073/pnas.1009945107.

681. Huh, M. S. *et al.* Compromised genomic integrity impedes muscle growth after Atrx inactivation. *J. Clin. Invest.* **122**, 4412–4423 (2012) DOI:10.1172/JCI63765.

682. Watson, L. A. *et al.* Atrx deficiency induces telomere dysfunction, endocrine defects, and reduced life span. *J. Clin. Invest.* **123**, 2049–2063 (2013) DOI:10.1172/JCI65634.

683. Kalb, R., Mallery, D. L., Larkin, C., Huang, J. T. J. & Hiom, K. BRCA1 is a histone-H2A-specific ubiquitin ligase. *Cell Rep.* **8**, 999–1005 (2014) DOI:10.1016/j.celrep.2014.07.025.

684. Zhu, Q. *et al.* BRCA1 tumour suppression occurs via heterochromatin-mediated silencing. *Nature* **477**, 179–84 (2011) DOI:10.1038/nature10371.

685. Thakar, A., Parvin, J. D. & Zlatanova, J. Brca1/bard1 e3 ubiquitin ligase can modify histones h2a and h2b in the nucleosome particle. *J. Biomol. Struct. Dyn.* **27**, 399–405 (2010) DOI:10.1080/07391102.2010.10507326.

686. Dellino, G. I. *et al.* Polycomb silencing blocks transcription initiation. *Mol. Cell* **13**, 887–893 (2004) DOI:10.1016/S1097-2765(04)00128-5.

687. Wang, H. *et al.* Role of histone H2A ubiquitination in Polycomb silencing. *Nature* **431**, 873–878 (2004) DOI:10.1038/nature02985.

688. Kleer, C. G. *et al.* EZH2 is a marker of aggressive breast cancer and promotes neoplastic transformation of breast epithelial cells. *Proc. Natl. Acad. Sci. U. S. A.* **100**, 11606–11611 (2003) DOI:10.1073/pnas.1933744100.

689. Lassman, A. B., Dai, C., Fuller, G. N. & Holland, E. C. Overexpression of c-MYC promotes an undifferentiated phenotype in cultured astrocytes and allows elevated Ras and Akt signaling to induce gliomas from GFAP-expressing cells in mice. *Neuron Glia Biol.* **1**, 157–163 (2004) DOI:10.1017/S1740925X04000249.

690. Orian, J. M. *et al.* Overexpression of multiple oncogenes related to histological grade of astrocytic glioma. *Br. J. Cancer* **66**, 106–112 (1992) DOI:10.1038/bjc.1992.225.

691. Valk-Lingbeek, M. E., Bruggeman, S. W. M. & Van Lohuizen, M. Stem cells and cancer: The polycomb connection. *Cell* (2004) doi:10.1016/j.cell.2004.08.005 DOI:10.1016/j.cell.2004.08.005.

692. Varambally, S. *et al.* The polycomb group protein EZH2 is involved in progression of prostate cancer. *Nature* **419**, 624–629 (2002) DOI:10.1038/nature01075.

693. Visser, H. P. J. *et al.* The Polycomb group protein EZH2 is upregulated in proliferating, cultured human mantle cell lymphoma. *Br. J. Haematol.* **112**, 950–958 (2001) DOI:10.1046/j.1365-2141.2001.02641.x.

694. Bruggeman, S. W. M. *et al.* Ink4a and Arf differentially affect cell proliferation and neural stem cell self-renewal in Bmi1-deficient mice. *Genes Dev.* **19**, 1438–1443 (2005) DOI:10.1101/gad.1299305.

695. Molofsky, A. V. *et al.* Bmi-1 dependence distinguishes neural stem cell self-renewal from progenitor proliferation. *Nature* **425**, 962–967 (2003) DOI:10.1038/nature02060.

696. Sharpless, N. E. & Depinho, R. A. The INK4A/ARF locus and its two gene products. *Curr. Opin. Genet. Dev.* **9**, 22–30 (1999) DOI:10.1016/S0959-437X(99)80004-5.

697. Ismail, H., Andrin, C., McDonald, D. & Hendzel, M. J. BMI1-mediated histone ubiquitylation promotes DNA double-strand break repair. *J. Cell Biol.* **191**, 45–60 (2010) DOI:10.1083/jcb.201003034.

698. Chagraoui, J. *et al.* An anticlastogenic function for the Polycomb Group gene Bmi1. *Proc. Natl. Acad. Sci.* **108**, 5284–5289 (2011) DOI:10.1073/pnas.1014263108.

699. Liu, J. *et al.* Bmi1 regulates mitochondrial function and the DNA damage response pathway. *Nature* **459**, 387–392 (2009) DOI:10.1038/nature08040.

700. Hernandez-Munoz, I., Taghavi, P., Kuijl, C., Neefjes, J. & van Lohuizen, M. Association of BMI1 with Polycomb Bodies Is Dynamic and Requires PRC2/EZH2 and the Maintenance DNA Methyltransferase DNMT1. *Mol. Cell. Biol.* **25**, 11047–11058 (2005) DOI:10.1128/mcb.25.24.11047-11058.2005.

701. Saurin, A. J. *et al.* The human polycomb group complex associates with pericentromeric heterochromatin to form a novel nuclear domain. *J. Cell Biol.* **142**, 887–898 (1998) DOI:10.1083/jcb.142.4.887.

702. Šmigová, J., Juda, P., Cmarko, D. & Raška, I. Fine structure of the 'pcG body' in human U-2 OS cells established by correlative light-electron microscopy. *Nucleus* **2**, 219–228 (2011) DOI:10.4161/nucl.2.3.15737.

703. Leeb, M. *et al.* Polycomb complexes act redundantly to repress genomic repeats and genes. *Genes Dev.* **24**, 265–276 (2010) DOI:10.1101/gad.544410.

704. Puschendorf, M. *et al.* PRC1 and Suv39h specify parental asymmetry at constitutive heterochromatin in early mouse embryos. *Nat. Genet.* **40**, 411–20 (2008) DOI:10.1038/ng.99.

705. Saksouk, N. *et al.* Redundant Mechanisms to Form Silent Chromatin at Pericentromeric Regions Rely on BEND3 and DNA Methylation. *Mol. Cell* **2**, 580–594 (2014) DOI:10.1016/j.molcel.2014.10.001.

706. Gargiulo, G. *et al.* InVivo RNAi Screen for BMI1 Targets Identifies TGF-β/BMP-ER Stress Pathways as Key Regulators of Neural- and Malignant Glioma-Stem Cell Homeostasis. *Cancer Cell* **23**, 660–676 (2013) DOI:10.1016/j.ccr.2013.03.030.

707. Tadeu, A. M. B. *et al.* CENP-V is required for centromere organization, chromosome alignment and cytokinesis. *EMBO J.* **27**, 2510–2522 (2008) DOI:10.1038/emboj.2008.175.

708. Shumaker, D. K. *et al.* Mutant nuclear lamin A leads to progressive alterations of epigenetic control in premature aging. *Proc. Natl. Acad. Sci.* **103**, 8703–8708 (2006) DOI:10.1073/pnas.0602569103.

709. Hwang, W. L., Deindl, S., Harada, B. T. & Zhuang, X. Histone H4 tail mediates allosteric regulation of nucleosome remodelling by linker DNA. *Nature* **512**, 213–217 (2014) DOI:10.1038/nature13380.

710. Hewish, D. R. & Burgoyne, L. A. Chromatin sub-structure. The digestion of chromatin DNA at regularly spaced sites by a nuclear deoxyribonuclease. *Biochem. Biophys. Res. Commun.* **52**, 504–510 (1973) DOI:10.1016/0006-291X(73)90740-7.

711. Kornberg, R. D., LaPointe, J. W. & Lorch, Y. Preparation of nucleosomes and chromatin. *Methods Enzymol.* **170**, 3–14 (1989) DOI:10.1016/0076-6879(89)70039-2.

712. Wu, C. The 5’ ends of drosophila heat shock genes in chromatin are hypersensitive to DNase I. *Nature* **286**, 854–860 (1980) DOI:10.1038/286854a0.

713. Román-Trufero, M. *et al.* Maintenance of undifferentiated state and self-renewal of embryonic neural stem cells by polycomb protein Ring1B. *Stem Cells* **27**, 1559–1570 (2009)

DOI:10.1002/stem.82.

714. Francis, N. J., Kingston, R. E. & Woodcock, C. L. Chromatin compaction by a polycomb group protein complex. *Science (80-.).* **306**, 1574–1577 (2004) DOI:10.1126/science.1100576.

715. Meshorer, E. *et al.* Hyperdynamic plasticity of chromatin proteins in pluripotent embryonic stem cells. *Dev. Cell* **10**, 105–116 (2006) DOI:10.1016/j.devcel.2005.10.017.

716. Varley, K. E. *et al.* Dynamic DNA methylation across diverse human cell lines and tissues. *Genome Res.* **23**, 555–567 (2013) DOI:10.1101/gr.147942.112.

717. Kohli, R. M. & Zhang, Y. TET enzymes, TDG and the dynamics of DNA demethylation. *Nature* **502**, 472–479 (2013) DOI:10.1038/nature12750.

718. Oberdoerffer, P. *et al.* SIRT1 Redistribution on Chromatin Promotes Genomic Stability but Alters Gene Expression during Aging. *Cell* **135**, 907–918 (2008) DOI:10.1016/j.cell.2008.10.025.

719. Larson, K. *et al.* Heterochromatin Formation Promotes Longevity and Represses Ribosomal RNA Synthesis. *PLoS Genet.* **8**, e1002473 (2012) DOI:10.1371/journal.pgen.1002473.

720. Pegoraro, G. *et al.* Ageing-related chromatin defects through loss of the NURD complex. *Nat. Cell Biol.* **11**, 1261–1267 (2009) DOI:10.1038/ncb1971.

721. Peng, J. C. & Karpen, G. H. Heterochromatic genome stability requires regulators of histone H3 K9 methylation. *PLoS Genet.* **5**, (2009) DOI:10.1371/journal.pgen.1000435.

722. Kim, J. *et al.* The n-SET Domain of Set1 Regulates H2B Ubiquitylation-Dependent H3K4 Methylation. *Mol. Cell* **49**, 1121–1133 (2013) DOI:10.1016/j.molcel.2013.01.034.

723. Wu, L. *et al.* ASH2L regulates ubiquitylation signaling to MLL: Trans-Regulation of H3 K4 methylation in higher eukaryotes. *Mol. Cell* **49**, 1108–1120 (2013) DOI:10.1016/j.molcel.2013.01.033.

724. McGinty, R. K., Henrici, R. C. & Tan, S. Crystal structure of the PRC1 ubiquitylation module bound to the nucleosome. *Nature* **514**, 591–596 (2014) DOI:10.1038/nature13890.

725. Satijn, D. P. *et al.* RING1 is associated with the polycomb group protein complex and acts as a transcriptional repressor. *Mol. Cell. Biol.* (2015) doi:10.1128/mcb.17.7.4105 DOI:10.1128/mcb.17.7.4105.

726. Kundu, S. *et al.* Polycomb Repressive Complex 1 Generates Discrete Compacted Domains that Change during Differentiation. *Mol. Cell* **65**, 432-446.e5 (2017) DOI:10.1016/j.molcel.2017.01.009.

727. Endoh, M. *et al.* PCGF6-PRC1 suppresses premature differentiation of mouse embryonic stem cells by regulating germ cell-related genes. *Elife* **6**, (2017) DOI:10.7554/eLife.21064.

728. Ito, T., Teo, Y. V., Evans, S. A., Neretti, N. & Sedivy Correspondence, J. M. Regulation of Cellular Senescence by Polycomb Chromatin Modifiers through Distinct DNA Damage- and Histone Methylation-Dependent Pathways. *Cell Rep.* **22**, 3480–3492 (2018) DOI:10.1016/j.celrep.2018.03.002.

729. Bracken, A. P. *et al.* The Polycomb group proteins bind throughout the INK4A-ARF locus and are disassociated in senescent cells. *Genes Dev.* **21**, 525–530 (2007) DOI:10.1101/gad.415507.

730. Dietrich, N. *et al.* Bypass of senescence by the polycomb group protein CBX8 through direct binding to the INK4A-ARF locus. *EMBO J.* **26**, 1637–48 (2007) DOI:10.1038/sj.emboj.7601632.

731. Abdouh, M., Hanna, R., El Hajjar, J., Flamier, A. & Bernier, G. The Polycomb Repressive Complex 1 Protein BMI1 Is Required for Constitutive Heterochromatin Formation and Silencing in Mammalian Somatic Cells. *J. Biol. Chem.* **291**, 182–197 (2016) DOI:10.1074/jbc.M115.662403.

732. Sun, W., Samimi, H., Gamez, M., Zare, H. & Frost, B. Pathogenic tau-induced piRNA depletion promotes neuronal death through transposable element dysregulation in neurodegenerative tauopathies. *Nat. Neurosci.* **21**, 1038–1048 (2018) DOI:10.1038/s41593-018-0194-1.

733. Guo, C. *et al.* Tau Activates Transposable Elements in Alzheimer's Disease. *Cell Rep.* **23**,

2874–2880 (2018) DOI:10.1016/j.celrep.2018.05.004.

734. Xia, X., Jiang, Q., McDermott, J. & Han, J.-D. J. Aging and Alzheimer's disease: Comparison and associations from molecular to system level. *Aging Cell* **17**, e12802 (2018) DOI:10.1111/acel.12802.

735. Guerreiro, R. & Bras, J. The age factor in Alzheimer's disease. *Genome Med.* **7**, 106 (2015) DOI:10.1186/S13073-015-0232-5.

736. Boccardi, V., Pelini, L., Ercolani, S., Ruggiero, C. & Mecocci, P. From cellular senescence to Alzheimer's disease: The role of telomere shortening. *Ageing Research Reviews* (2015) doi:10.1016/j.arr.2015.04.003 DOI:10.1016/j.arr.2015.04.003.

737. Reddy, P. H. *et al.* MicroRNAs, Aging, Cellular Senescence, and Alzheimer's Disease. in *Progress in Molecular Biology and Translational Science* (2017). doi:10.1016/bs.pmbts.2016.12.009 DOI:10.1016/bs.pmbts.2016.12.009.

738. Frost, B., Bardai, F. H. H. & Feany, M. B. B. Lamin Dysfunction Mediates Neurodegeneration in Tauopathies. *Curr. Biol.* **26**, 129–136 (2016) DOI:10.1016/j.cub.2015.11.039.

739. Cheung, I., Schertzer, M., Rose, A. & Lansdorp, P. M. Disruption of dog-1 in Caenorhabditis elegans triggers deletions upstream of guanine-rich DNA. *Nat. Genet.* (2002) doi:10.1038/ng928 DOI:10.1038/ng928.

740. Maizels, N. & Gray, L. T. The G4 Genome. *PLoS Genet.* **9**, e1003468 (2013) DOI:10.1371/journal.pgen.1003468.

741. Hänsel-Hertsch, R. *et al.* G-quadruplex structures mark human regulatory chromatin. *Nat. Genet.* (2016) doi:10.1038/ng.3662 DOI:10.1038/ng.3662.

742. Johnson, J. E., Cao, K., Ryvkin, P., Wang, L. S. & Johnson, F. B. Altered gene expression in the Werner and Bloom syndromes is associated with sequences having G-quadruplex forming potential. *Nucleic Acids Res.* (2009) doi:10.1093/nar/gkp1103 DOI:10.1093/nar/gkp1103.

743. Sidorova, J. M. Roles of the Werner syndrome RecQ helicase in DNA replication. *DNA Repair*

(2008) doi:10.1016/j.dnarep.2008.07.017 DOI:10.1016/j.dnarep.2008.07.017.

744. Rogakou, E. P., Pilch, D. R., Orr, A. H., Ivanova, V. S. & Bonner, W. M. DNA double-stranded breaks induce histone H2AX phosphorylation on serine 139. *J. Biol. Chem.* (1998) doi:10.1074/jbc.273.10.5858 DOI:10.1074/jbc.273.10.5858.

745. Bhaskara, S. *et al.* Hdac3 is essential for the maintenance of chromatin structure and genome stability. *Cancer Cell* (2010) doi:10.1016/j.ccr.2010.10.022 DOI:10.1016/j.ccr.2010.10.022.

746. Blander, G. & Guarente, L. The Sir2 Family of Protein Deacetylases. *Annu. Rev. Biochem.* **73**, 417–435 (2004) DOI:10.1146/annurev.biochem.73.011303.073651.

747. Gray, R. D., Trent, J. O. & Chaires, J. B. Folding and unfolding pathways of the human telomeric G-quadruplex. *J. Mol. Biol.* (2014) doi:10.1016/j.jmb.2014.01.009 DOI:10.1016/j.jmb.2014.01.009.

748. Panier, S. & Boulton, S. J. Double-strand break repair: 53BP1 comes into focus. *Nature Reviews Molecular Cell Biology* (2014) doi:10.1038/nrm3719 DOI:10.1038/nrm3719.

749. Leemans, C. *et al.* Promoter-Intrinsic and Local Chromatin Features Determine Gene Repression in LADs. *Cell* **177**, 852-864.e14 (2019) DOI:10.1016/j.cell.2019.03.009.

750. Vazquez, B. N. *et al.* SIRT7 mediates L1 elements transcriptional repression and their association with the nuclear lamina. *Nucleic Acids Res.* (2019) doi:10.1093/nar/gkz519 DOI:10.1093/nar/gkz519.

751. Hernández, F., García-García, E. & Avila, J. Microtubule Depolymerization and Tau Phosphorylation. *J. Alzheimer's Dis.* **37**, 507–513 (2013) DOI:10.3233/JAD-130545.

752. Hernandez, F., Lucas, J. J. & Avila, J. GSK3 and tau: two convergence points in Alzheimer's disease. *J. Alzheimers. Dis.* **33 Suppl 1**, S141-4 (2013) DOI:10.3233/JAD-2012-129025.

753. Mungamuri, S. K. *et al.* p53-mediated heterochromatin reorganization regulates its cell fate decisions. *Nat. Struct. Mol. Biol.* **19**, 478–84, S1 (2012) DOI:10.1038/nsmb.2271.

754. Meuleman, W. *et al.* Constitutive nuclear lamina-genome interactions are highly conserved and associated with A/T-rich sequence. *Genome Res.* **23**, 270–280 (2013) DOI:10.1101/gr.141028.112.

755. Varizhuk, A. *et al.* The expanding repertoire of G4 DNA structures. *Biochimie* (2017) doi:10.1016/j.biochi.2017.01.003 DOI:10.1016/j.biochi.2017.01.003.

756. Piazza, A. *et al.* Non-Canonical G-quadruplexes cause the hCEB1 minisatellite instability in Saccharomyces cerevisiae. *Elife* (2017) doi:10.7554/elife.26884 DOI:10.7554/elife.26884.

757. Kikin, O., D'Antonio, L. & Bagga, P. S. QGRS Mapper: A web-based server for predicting G-quadruplexes in nucleotide sequences. *Nucleic Acids Res.* **34**, W676-82 (2006) DOI:10.1093/nar/gkl253.

758. D'Antonio, L. & Bagga, P. Computational methods for predicting intramolecular G-quadruplexes in nucleotide sequences. in *Proceedings - 2004 IEEE Computational Systems Bioinformatics Conference, CSB 2004* 590–591 (2004). doi:10.1109/csb.2004.1332508 DOI:10.1109/csb.2004.1332508.

759. Fiszbein, A., Krick, K. S., Begg, B. E. & Burge, C. B. Exon-Mediated Activation of Transcription Starts. *Cell* **179**, 1551-1565.e17 (2019) DOI:10.1016/j.cell.2019.11.002.

760. Levy, M. A., Kernohan, K. D., Jiang, Y. & Bérubé, N. G. ATRX promotes gene expression by facilitating transcriptional elongation through guanine-rich coding regions. *Hum. Mol. Genet.* **24**, 1–12 (2014) DOI:10.1093/hmg/ddu596.

761. Murat, P. & Balasubramanian, S. Existence and consequences of G-quadruplex structures in DNA. *Current Opinion in Genetics and Development* (2014) doi:10.1016/j.gde.2013.10.012 DOI:10.1016/j.gde.2013.10.012.

762. van Steensel, B. & Belmont, A. S. Lamina-Associated Domains: Links with Chromosome Architecture, Heterochromatin, and Gene Repression. *Cell* vol. 169 780–791 (2017) DOI:10.1016/j.cell.2017.04.022.

763. Brouha, B. *et al.* Hot L1s account for the bulk of retrotransposition in the human

population. *Proc. Natl. Acad. Sci.* **100**, 5280–5285 (2003) DOI:10.1073/pnas.0831042100.

764. Flasch, D. A. *et al.* Genome-wide de novo L1 Retrotransposition Connects Endonuclease Activity with Replication. *Cell* **177**, 837-851.e28 (2019) DOI:10.1016/j.cell.2019.02.050.

765. Howell, R. & Usdin, K. The ability to form intrastrand tetraplexes is an evolutionarily conserved feature of the 3' end of L1 retrotransposons. *Mol. Biol. Evol.* **14**, 144–155 (1997) DOI:10.1093/oxfordjournals.molbev.a025747.

766. Raj, T. *et al.* Integrative transcriptome analyses of the aging brain implicate altered splicing in Alzheimer's disease susceptibility. *Nat. Genet.* **50**, 1584–1592 (2018) DOI:10.1038/s41588-018-0238-1.

767. Llorens-MartÃ-n, M. *et al.* Selective alterations of neurons and circuits related to early memory loss in Alzheimerâ€™s disease. *Front. Neuroanat.* (2014) doi:10.3389/fnana.2014.00038 DOI:10.3389/fnana.2014.00038.

768. Arendt, T. *et al.* Plastic Neuronal Remodeling Is Impaired in Patients with Alzheimer's Disease Carrying Apolipoprotein ε4 Allele. *J. Neurosci.* (2018) doi:10.1523/jneurosci.17-02-00516.1997 DOI:10.1523/jneurosci.17-02-00516.1997.

769. Skaper, S. D., Facci, L., Zusso, M. & Giusti, P. Synaptic Plasticity, Dementia and Alzheimer Disease. *CNS Neurol. Disord. - Drug Targets* (2017) doi:10.2174/1871527316666170113120853 DOI:10.2174/1871527316666170113120853.

770. Uchida, Y. *et al.* Semaphorin3A signalling is mediated via sequential Cdk5 and GSK3β phosphorylation of CRMP2: Implication of common phosphorylating mechanism underlying axon guidance and Alzheimer's disease. *Genes to Cells* (2005) doi:10.1111/j.1365-2443.2005.00827.x DOI:10.1111/j.1365-2443.2005.00827.x.

771. Good, P. F. *et al.* A role for semaphorin 3A signaling in the degeneration of hippocampal neurons during Alzheimer's disease. *J. Neurochem.* (2004) doi:10.1111/j.1471-4159.2004.02766.x DOI:10.1111/j.1471-4159.2004.02766.x.

772. Van Battum, E. Y., Brignani, S. & Pasterkamp, R. J. Axon guidance proteins in neurological

disorders. *The Lancet Neurology* (2015) doi:10.1016/S1474-4422(14)70257-1 DOI:10.1016/S1474-4422(14)70257-1.

773. Lerdkrai, C. *et al.* Intracellular Ca 2+ stores control in vivo neuronal hyperactivity in a mouse model of Alzheimer's disease . *Proc. Natl. Acad. Sci.* (2018) doi:10.1073/pnas.1714409115 DOI:10.1073/pnas.1714409115.

774. Toni, N. & Sultan, S. Synapse formation on adult-born hippocampal neurons. *Eur. J. Neurosci.* (2011) doi:10.1111/j.1460-9568.2011.07604.x DOI:10.1111/j.1460-9568.2011.07604.x.

775. Zhang, Y.-J. *et al.* Heterochromatin anomalies and double-stranded RNA accumulation underlie C9orf72 poly(PR) toxicity. *Science (80-.).* **363**, eaav2606 (2019) DOI:10.1126/SCIENCE.AAV2606.

776. Liu, E. Y. *et al.* Loss of Nuclear TDP-43 Is Associated with Decondensation of LINE Retrotransposons. *Cell Rep.* **27**, 1409-1421.e6 (2019) DOI:10.1016/j.celrep.2019.04.003.

777. Belzil, V. V. *et al.* Reduced C9orf72 gene expression in c9FTD/ALS is caused by histone trimethylation, an epigenetic event detectable in blood. *Acta Neuropathol.* **126**, 895–905 (2013) DOI:10.1007/s00401-013-1199-1.

778. Krug, L. *et al.* Retrotransposon activation contributes to neurodegeneration in a Drosophila TDP-43 model of ALS. *PLoS Genet.* **13**, e1006635 (2017) DOI:10.1371/journal.pgen.1006635.

779. Meng, S. *et al.* Identification and characterization of Bmi-1-responding element within the human p16 promoter. *J. Biol. Chem.* (2010) doi:10.1074/jbc.M110.133686 DOI:10.1074/jbc.M110.133686.

780. Lerdrup, M., Johansen, J. V., Agrawal-Singh, S. & Hansen, K. An interactive environment for agile analysis and visualization of ChIP-sequencing data. *Nat. Struct. Mol. Biol.* (2016) doi:10.1038/nsmb.3180 DOI:10.1038/nsmb.3180.

781. Gel, B. *et al.* RegioneR: An R/Bioconductor package for the association analysis of genomic

regions based on permutation tests. *Bioinformatics* (2016) doi:10.1093/bioinformatics/btv562 DOI:10.1093/bioinformatics/btv562.

782. Karolchik, D. *et al.* The UCSC Genome Browser Database. *Nucleic Acids Res.* (2003) doi:10.1093/nar/gkg129 DOI:10.1093/nar/gkg129.

783. Bailey, T. L. *et al.* MEME Suite: Tools for motif discovery and searching. *Nucleic Acids Res.* (2009) doi:10.1093/nar/gkp335 DOI:10.1093/nar/gkp335.

784. Machanick, P. & Bailey, T. L. MEME-ChIP: Motif analysis of large DNA datasets. *Bioinformatics* (2011) doi:10.1093/bioinformatics/btr189 DOI:10.1093/bioinformatics/btr189.

785. Ashburner, M. *et al.* Gene ontology: Tool for the unification of biology. *Nature Genetics* (2000) doi:10.1038/75556 DOI:10.1038/75556.

786. Carbon, S. *et al.* The Gene Ontology Resource: 20 years and still GOing strong. *Nucleic Acids Res.* **47**, D330–D338 (2019) DOI:10.1093/nar/gky1055.

787. Mi, H. *et al.* PANTHER version 11: Expanded annotation data from Gene Ontology and Reactome pathways, and data analysis tool enhancements. *Nucleic Acids Res.* (2017) doi:10.1093/nar/gkw1138 DOI:10.1093/nar/gkw1138.

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

789. Team, R. C. The R Project for Statistical Computing. *Http://Www.R-Project.Org/* 1–12 https://www.r-project.org/ (2013).

790. Mendel, G. Experiments in plant Hybridization. *Verhandlungen des naturforschenden Vereines in Brünn* http://www.esp.org/foundations/genetics/classical/gm-65.pdf (1866).

791. FRANKLIN, R. E. & GOSLING, R. G. Molecular Configuration in Sodium Thymonucleate. *Nature* **171**, 740–741 (1953) DOI:10.1038/171740a0.

792. Nirenberg, M. W. & Matthaei, J. H. The dependence of cell-free protein synthesis in E. coli

upon naturally occurring or synthetic polyribonucleotides. *Proc. Natl. Acad. Sci.* **47**, 1588–1602 (1961) DOI:10.1073/pnas.47.10.1588.

793. DeLisi, C. Meetings that changed the world: Santa Fe 1986: Human genome baby-steps. *Nature* **455**, 876–7 (2008) DOI:10.1038/455876a.

794. Delisi, Charles. The Human Genome Project. *Am. Sci.* **76**, (1988).

795. DeLisi, C. Genomes: 15 Years Later. *Human Genome News* vol. 11 5–6 https://web.archive.org/web/20040904183906/http://genome.gsc.riken.go.jp:80/hgmis/publicat/hgn/v11n3/05delisi.html (2001).

796. Holliday, R. DNA Methylation and Epigenetic Inheritance. *Philos. Trans. R. Soc. B Biol. Sci.* **326**, 329–338 (1990) DOI:10.1098/rstb.1990.0015.

797. Jeltsch, A., Walter, J., Reinhardt, R. & Platzer, M. German human methylome project started. *Cancer Res.* **66**, 7378 (2006) DOI:10.1158/0008-5472.CAN-06-1071.

798. Romanoski, C. E., Glass, C. K., Stunnenberg, H. G., Wilson, L. & Almouzni, G. Epigenomics: Roadmap for regulation. *Nature* **518**, 314–6 (2015) DOI:10.1038/518314a.

799. Dong, X. & Weng, Z. The correlation between histone modifications and gene expression. *Epigenomics* **5**, 113–6 (2013) DOI:10.2217/epi.13.13.

800. Zhu, Q. & Wani, A. A. Histone modifications: crucial elements for damage response and chromatin restoration. *J. Cell. Physiol.* **223**, 283–8 (2010) DOI:10.1002/jcp.22060.

801. Wilson, W. D. & Sugiyama, H. First International Meeting on Quadruplex DNA. *ACS Chem. Biol.* **2**, 589–94 (2007) DOI:10.1021/cb7001686.

802. Arthanari, H. & Bolton, P. H. Functional and dysfunctional roles of quadruplex DNA in cells. *Chem. Biol.* **8**, 221–30 (2001) DOI:10.1016/S1074-5521(01)00007-2.

803. Wu, Y. & Brosh, R. M. G-quadruplex nucleic acids and human disease. *FEBS J.* **277**, 3470–88 (2010) DOI:10.1111/j.1742-4658.2010.07760.x.

804. Zhou, G. *et al.* Telomere targeting with a novel G-quadruplex-interactive ligand BRACO-19

induces T-loop disassembly and telomerase displacement in human glioblastoma cells. *Oncotarget* (2016) doi:10.18632/oncotarget.7483 DOI:10.18632/oncotarget.7483.

805. Kamath-Loeb, A. S., Loeb, L. A., Johansson, E., Burgers, P. M. & Fry, M. Interactions between the Werner syndrome helicase and DNA polymerase delta specifically facilitate copying of tetraplex and hairpin structures of the d(CGG)n trinucleotide repeat sequence. *J. Biol. Chem.* **276**, 16439–46 (2001) DOI:10.1074/jbc.M100253200.

806. Pirzio, L. M., Pichierri, P., Bignami, M. & Franchitto, A. Werner syndrome helicase activity is essential in maintaining fragile site stability. *J. Cell Biol.* **180**, 305–14 (2008) DOI:10.1083/jcb.200705126.

807. Piazza, A. *et al.* Short loop length and high thermal stability determine genomic instability induced by G-quadruplex-forming minisatellites. *EMBO J.* **34**, e201490702 (2015) DOI:10.15252/embj.201490702.

808. Maity, A., Winnerdy, F. R., Chang, W. D., Chen, G. & Phan, A. T. Intra-locked G-quadruplex structures formed by irregular DNA G-rich motifs. *Nucleic Acids Res.* (2020) doi:10.1093/nar/gkaa008 DOI:10.1093/nar/gkaa008.

809. Beniaminov, A., Shchyolkina, A. & Kaluzhny, D. Conformational features of intramolecular G4-DNA constrained by single-nucleotide loops. *Biochimie* (2019) doi:10.1016/j.biochi.2019.02.013 DOI:10.1016/j.biochi.2019.02.013.

810. Dingley, A. J., Peterson, R. D., Grzesiek, S. & Feigon, J. Characterization of the cation and temperature dependence of DNA quadruplex hydrogen bond properties using high-resolution NMR. *J. Am. Chem. Soc.* (2005) doi:10.1021/ja0540369 DOI:10.1021/ja0540369.

811. Hao, F., Ma, Y. & Guan, Y. Effects of central loop length and metal ions on the thermal stability of G-quadruplexes. *Molecules* (2019) doi:10.3390/molecules24101863 DOI:10.3390/molecules24101863.

812. Cheng, M. *et al.* Loop permutation affects the topology and stability of G-quadruplexes. *Nucleic Acids Res.* **46**, 9264–9275 (2018) DOI:10.1093/nar/gky757.

813. Stadlbauer, P. *et al.* Parallel G-triplexes and G-hairpins as potential transitory ensembles in the folding of parallel-stranded DNA G-Quadruplexes. *Nucleic Acids Res.* (2019) doi:10.1093/nar/gkz610 DOI:10.1093/nar/gkz610.

814. Marchand, A. & Gabelica, V. Folding and misfolding pathways of G-quadruplex DNA. *Nucleic Acids Res.* (2016) doi:10.1093/nar/gkw970 DOI:10.1093/nar/gkw970.

815. Lee, J. Y., Okumus, B., Kim, D. S. & Ha, T. Extreme conformational diversity in human telomeric DNA. *Proc. Natl. Acad. Sci. U. S. A.* (2005) doi:10.1073/pnas.0506144102 DOI:10.1073/pnas.0506144102.

816. Tippana, R., Xiao, W. & Myong, S. G-quadruplex conformation and dynamics are determined by loop length and sequence. *Nucleic Acids Res.* (2014) doi:10.1093/nar/gku464 DOI:10.1093/nar/gku464.

817. Zubradt, M. *et al.* DMS-MaPseq for genome-wide or targeted RNA structure probing in vivo. *Nat. Methods* **14**, 75–82 (2016) DOI:10.1038/nmeth.4057.

818. Havrila, M. *et al.* Structural dynamics of propeller loop: Towards folding of RNA G-quadruplex. *Nucleic Acids Res.* (2018) doi:10.1093/nar/gky712 DOI:10.1093/nar/gky712.

819. Yang, S. Y. *et al.* Transcriptome-wide identification of transient RNA G-quadruplexes in human cells. *Nat. Commun.* **9**, (2018) DOI:10.1038/s41467-018-07224-8.

820. Cammas, A. & Millevoi, S. RNA G-quadruplexes: emerging mechanisms in disease. *Nucleic Acids Res.* (2017) doi:10.1093/nar/gkw1280 DOI:10.1093/nar/gkw1280.

821. Fay, M. M., Lyons, S. M. & Ivanov, P. RNA G-Quadruplexes in Biology: Principles and Molecular Mechanisms. *Journal of Molecular Biology* vol. 429 2127–2147 (2017) DOI:10.1016/j.jmb.2017.05.017.

822. Dmitrijeva, M., Ossowski, S., Serrano, L. & Schaefer, M. H. Tissue-specific DNA methylation loss during ageing and carcinogenesis is linked to chromosome structure, replication timing and cell division rates. *Nucleic Acids Res.* (2018) doi:10.1093/nar/gky498 DOI:10.1093/nar/gky498.

823. Pegoraro, G. & Misteli, T. The central role of chromatin maintenance in aging. *Aging (Albany. NY).* **1**, 1017–1022 (2009) DOI:10.18632/aging.100106.

824. Benayoun, B. A., Pollina, E. A. & Brunet, A. Epigenetic regulation of ageing: Linking environmental inputs to genomic stability. *Nature Reviews Molecular Cell Biology* vol. 16 593–610 (2015) DOI:10.1038/nrm4048.

825. Epel, E. S. *et al.* Accelerated telomere shortening in response to life stress. *Proc. Natl. Acad. Sci. U. S. A.* **101**, 17312–17315 (2004) DOI:10.1073/pnas.0407162101.

826. Horvath, S. *et al.* Obesity accelerates epigenetic aging of human liver. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 15538–15543 (2014) DOI:10.1073/pnas.1412759111.

827. Ciccarone, F., Tagliatesta, S., Caiafa, P. & Zampieri, M. DNA methylation dynamics in aging: how far are we from understanding the mechanisms? *Mechanisms of Ageing and Development* vol. 174 3–17 (2018) DOI:10.1016/j.mad.2017.12.002.

828. Tsygankov, D., Liu, Y., Sanoff, H. K., Sharpless, N. E. & Elston, T. C. A quantitative model for age-dependent expression of the p16INK4a tumor suppressor. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 16562–16567 (2009) DOI:10.1073/pnas.0904405106.

829. Hannum, G. *et al.* Genome-wide Methylation Profiles Reveal Quantitative Views of Human Aging Rates. *Mol. Cell* **49**, 359–367 (2013) DOI:10.1016/j.molcel.2012.10.016.

830. Pérez, R. F., Tejedor, J. R., Bayón, G. F., Fernández, A. F. & Fraga, M. F. Distinct chromatin signatures of DNA hypomethylation in aging and cancer. *Aging Cell* **17**, (2018) DOI:10.1111/acel.12744.

831. Horvath, S. DNA methylation age of human tissues and cell types. *Genome Biol.* **14**, R115 (2013) DOI:10.1186/gb-2013-14-10-r115.

832. Weidner, C. I. *et al.* Aging of blood can be tracked by DNA methylation changes at just three CpG sites. *Genome Biol.* **15**, (2014) DOI:10.1186/gb-2014-15-2-r24.

833. Bocklandt, S. *et al.* Epigenetic predictor of age. *PLoS One* **6**, (2011) DOI:10.1371/journal.pone.0014821.

834. Zou, J., Lippert, C., Heckerman, D., Aryee, M. & Listgarten, J. Epigenome-wide association studies without the need for cell-type composition. *Nat. Methods* **11**, 309–311 (2014) DOI:10.1038/nmeth.2815.

835. Kim, E. C. & Kim, J. R. Senotherapeutics: Emerging strategy for healthy aging and age-related disease. *BMB Reports* vol. 52 47–55 (2019) DOI:10.5483/BMBRep.2019.52.1.293.

836. Lardenoije, R., Pishva, E., Lunnon, K. & van den Hove, D. L. Neuroepigenetics of Aging and Age-Related Neurodegenerative Disorders. in *Progress in Molecular Biology and Translational Science* vol. 158 49–82 (Elsevier B.V., 2018). DOI:10.1016/bs.pmbts.2018.04.008.

837. Lardenoije, R. *et al.* The epigenetics of aging and neurodegeneration. *Progress in Neurobiology* vol. 131 21–64 (2015) DOI:10.1016/j.pneurobio.2015.05.002.

838. Park, M. *et al.* Age-associated chromatin relaxation is enhanced in Huntington's disease mice. *Aging (Albany. NY).* **9**, 803–822 (2017) DOI:10.18632/aging.101193.

839. Song, W., Zsindely, N., Faragó, A., Marsh, J. L. & Bodai, L. Systematic genetic interaction studies identify histone demethylase utx as potential target for ameliorating huntington's disease. *Hum. Mol. Genet.* (2018) doi:10.1093/hmg/ddx432 DOI:10.1093/hmg/ddx432.

840. Bassi, S., Tripathi, T., Monziani, A., Di Leva, F. & Biagioli, M. Epigenetics of huntington's disease. *Adv. Exp. Med. Biol.* **978**, 277–299 (2017) DOI:10.1007/978-3-319-53889-1_15.

841. Zsindely, N. & Bodai, L. Histone methylation in Huntington's disease: are bivalent promoters the critical targets? *Neural Regen. Res.* **13**, 1191–1192 (2018) DOI:10.4103/1673-5374.235029.

842. McClymont, S. A. *et al.* Parkinson-Associated SNCA Enhancer Variants Revealed by Open Chromatin in Mouse Dopamine Neurons. *Am. J. Hum. Genet.* **103**, 874–892 (2018) DOI:10.1016/j.ajhg.2018.10.018.

843. Wüllner, U., Kaut, O., deBoni, L., Piston, D. & Schmitt, I. DNA methylation in Parkinson's disease. *Journal of Neurochemistry* vol. 139 Suppl 1 108–120 (2016)

DOI:10.1111/jnc.13646.

844. Habibi, E., Masoudi-Nejad, A., Abdolmaleky, H. M. & Haggarty, S. J. Emerging roles of epigenetic mechanisms in Parkinson's disease. *Functional and Integrative Genomics* vol. 11 523–537 (2011) DOI:10.1007/s10142-011-0246-z.

845. Tibshirani, M. *et al.* Dysregulation of chromatin remodelling complexes in amyotrophic lateral sclerosis. *Hum. Mol. Genet.* (2017) doi:10.1093/hmg/ddx301 DOI:10.1093/hmg/ddx301.

846. Jimenez-Pacheco, A. *et al.* Epigenetic mechanisms of gene regulation in amyotrophic lateral sclerosis. in *Advances in Experimental Medicine and Biology* vol. 978 255–275 (Springer New York LLC, 2017). DOI:10.1007/978-3-319-53889-1_14.

847. Bennett, S. A., Tanaz, R., Cobos, S. N. & Torrente, M. P. Epigenetics in amyotrophic lateral sclerosis: a role for histone post-translational modifications in neurodegenerative disease. *Translational Research* vol. 204 19–30 (2019) DOI:10.1016/j.trsl.2018.10.002.

848. Masala, A. *et al.* Epigenetic Changes Associated with the Expression of Amyotrophic Lateral Sclerosis (ALS) Causing Genes. *Neuroscience* **390**, 1–11 (2018) DOI:10.1016/j.neuroscience.2018.08.009.

849. Cacabelos, R. Parkinson's disease: From pathogenesis to pharmacogenomics. *International Journal of Molecular Sciences* vol. 18 (2017) DOI:10.3390/ijms18030551.

850. Jakubowski, J. L. & Labrie, V. Epigenetic Biomarkers for Parkinson's Disease: From Diagnostics to Therapeutics. *Journal of Parkinson's Disease* vol. 7 1–12 (2017) DOI:10.3233/JPD-160914.

851. Ciceri, F., Rotllant, D. & Maes, T. Understanding Epigenetic Alterations in Alzheimer's and Parkinson's Disease: Towards Targeted Biomarkers and Therapies. *Curr. Pharm. Des.* **23**, 839–857 (2017) DOI:10.2174/1381612823666170124121140.

852. Irwin, M. H., Moos, W. H., Faller, D. V., Steliou, K. & Pinkert, C. A. Epigenetic Treatment of Neurodegenerative Disorders: Alzheimer and Parkinson Diseases. *Drug Dev. Res.* **77**, 109–

123 (2016) DOI:10.1002/ddr.21294.

853. De Magis, A. *et al.* DNA damage and genome instability by G-quadruplex ligands are mediated by R loops in human cancer cells. *Proc. Natl. Acad. Sci. U. S. A.* **116**, 816–825 (2019) DOI:10.1073/pnas.1810409116.

854. Moruno-Manchon, J. F. *et al.* The G-quadruplex DNA stabilizing drug pyridostatin promotes DNA damage and downregulates transcription of Brca1 in neurons. *Aging (Albany. NY).* **9**, 1957–1970 (2017) DOI:10.18632/aging.101282.

855. van Kregten, M. & Tijsterman, M. The repair of G-quadruplex-induced DNA damage. *Experimental Cell Research* vol. 329 178–183 (2014) DOI:10.1016/j.yexcr.2014.08.038.

856. Santos-Pereira, J. M. & Aguilera, A. R loops: New modulators of genome dynamics and function. *Nature Reviews Genetics* vol. 16 583–597 (2015) DOI:10.1038/nrg3961.

857. Q, A.-H. & Y, Y. R-loop: An Emerging Regulator of Chromatin Dynamics. *Acta Biochim. Biophys. Sin. (Shanghai).* **48**, (2016) DOI:10.1093/ABBS/GMW052.

858. Drolet, M. & Brochu, J. R-loop-dependent replication and genomic instability in bacteria. *DNA Repair* vol. 84 (2019) DOI:10.1016/j.dnarep.2019.102693.

859. Freudenreich, C. H. R-loops: targets for nuclease cleavage and repeat instability. *Current Genetics* vol. 64 789–794 (2018) DOI:10.1007/s00294-018-0806-z.

860. Aguilera, A. & García-Muse, T. R Loops: From Transcription Byproducts to Threats to Genome Stability. *Mol. Cell* **46**, 115–124 (2012) DOI:10.1016/j.molcel.2012.04.009.

861. Stirling, P. C. & Hieter, P. Canonical DNA Repair Pathways Influence R-Loop-Driven Genome Instability. *Journal of Molecular Biology* vol. 429 3132–3138 (2017) DOI:10.1016/j.jmb.2016.07.014.

862. Yang, D. & Okamoto, K. Structural insights into G-quadruplexes: Towards new anticancer drugs. *Future Medicinal Chemistry* vol. 2 619–646 (2010) DOI:10.4155/fmc.09.172.

863. Düchler, M. G-quadruplexes: targets and tools in anticancer drug design. *J. Drug Target.*

20, 389–400 (2012) DOI:10.3109/1061186X.2012.669384.

864. De Magis, A. *et al.* Zuo1 supports G4 structure formation and directs repair toward nucleotide excision repair. *Nat. Commun.* **11**, 1–12 (2020) DOI:10.1038/s41467-020-17701-8.

865. Dose-escalation Study of Quarfloxin in Patients With Advanced Solid Tumors or Lymphomas - Full Text View - ClinicalTrials.gov. https://clinicaltrials.gov/ct2/show/NCT00955292?term=quadruplex&draw=2&rank=2.

866. Lensing, S. V. *et al.* DSBCapture: In situ capture and sequencing of DNA breaks. *Nat. Methods* **13**, 855–857 (2016) DOI:10.1038/nmeth.3960.

867. Bouwman, B. A. M. & Crosetto, N. Endogenous DNA double-strand breaks during DNA transactions: Emerging insights and methods for genome-wide profiling. *Genes* vol. 9 (2018) DOI:10.3390/genes9120632.

868. Tauchi, T. *et al.* Activity of a novel G-quadruplex-interactive telomerase inhibitor, telomestatin (SOT-095), against human leukemia cells: involvement of ATM-dependent DNA damage response pathways. *Oncogene* **22**, 5338–5347 (2003) DOI:10.1038/sj.onc.1206833.

869. Tauchi, T. *et al.* Telomerase inhibition with a novel G-quadruplex-interactive agent, telomestatin: In vitro and in vivo studies in acute leukemia. *Oncogene* **25**, 5719–5725 (2006) DOI:10.1038/sj.onc.1209577.

870. Phatak, P. *et al.* Telomere uncapping by the G-quadruplex ligand RHPS4 inhibits clonogenic tumour cell growth in vitro and in vivo consistent with a cancer stem cell targeting mechanism. *Br. J. Cancer* **96**, 1223–1233 (2007) DOI:10.1038/sj.bjc.6603691.

871. Leonetti, C. *et al.* G-quadruplex ligand RHPS4 potentiates the antitumor activity of camptothecins in preclinical models of solid tumors. *Clin. Cancer Res.* **14**, 7284–7291 (2008) DOI:10.1158/1078-0432.CCR-08-0941.

872. Lixia, G., Fei, Y., Jiajia, J. & Jianhui, L. Triethylene tetramine, a novel ligand of G-quadruplex,

induces senescence of MCF-7 cells. *Biotechnol. Lett.* **30**, 47–53 (2008) DOI:10.1007/s10529-007-9513-4.

873. Shalaby, T. *et al.* Disabling c-Myc in childhood medulloblastoma and atypical teratoid/rhabdoid tumor cells by the potent G-quadruplex interactive agent S2T1-6OTD. *Mol. Cancer Ther.* **9**, 167–179 (2010) DOI:10.1158/1535-7163.MCT-09-0586.

6 Annex I: supplementary data related to chapter 3

Gene	Forward	Reverse
Hprt	5'-ACTGTAATGATCAGTCAACGGG-3'	5_-GGCCTGTATCCAACACTTGG-3'
Bmi1	5'-GGAGACCAGCAAGTATTGTCCTATTTG-3'	5'-CTTACGATGCCCAGCAGCAATG-3'
P16	5'-CAACGCCCCGAACTCTTTC-3'	5_-GCAGAAGAGCTGCTACGTGAAC-3'
Line	5′-TGGCTTGTGCTGTAAGATCG-3′	5′-TCTGTTGGTGGTCTTTTTGTC-3′
Sine	5′-GAGCACACCCATGCACATAC-3′	5′-AAAGGCATGCACCTCTACCACC-3′
Min. sat	5′-TTGGAAACGGGATTTGTAGA-3'	5′-CGGTTTCCAACATATGTGTTTT-3′
Maj. sat	5′-GGCGAGAAAACTGAAAATCACG-3′	5′-CTTGCCATATTCCACGTCCT-3′
IAP1	5′-CGCTCCGGTAGAATACTTAC-3′	5′-TGCCATGCCGGCGAGCCTGT-3′

Table 11 . Mouse RT-qPCR primers

Gene	Forward	Reverse
Line	5′-TGGCTTGTGCTGTAAGATCG-3′	5′-TCTGTTGGTGGTCTTTTTGTC-3′
Sine	5′-GAGCACACCCATGCACATAC-3′	5′-AAAGGCATGCACCTCTACCACC-3′
Min. sat	5′-TTGGAAACGGGATTTGTAGA-3’	5′-CGGTTTCCAACATATGTGTTTT-3′
Maj. sat	5′-GGCGAGAAAACTGAAAATCACG-3′	5′-CTTGCCATATTCCACGTCCT-3′
IAP1	5′-CGCTCCGGTAGAATACTTAC-3′	5′-TGCCATGCCGGCGAGCCTGT-3′
HoxA7.1	5’-GTGGGCAAAGAGTGGATTTC-3’	5’-CCCCGACAACCTCATACCTA-3’
Globin	5′-CAGTGAGTGGCACAGCATCC-3′	5′-CAGTCAGGTGCACCATGATGT-3′

Table 12 . Mouse ChIP-qPCR primers

Gene	Forward	Reverse
GAP-DH	5’-TCACCAGGGCTGCTTTTAAC-3’	5’-ATCCACAGTCTTCTGGGTGG-3’
BMI1	5’-AATCCCCACCTGATGTGTGT-3’	5’-GCTGGTCTCCAGGTAACGAA-3’
P16	5’-GGGTTTTCGTGGTTCACATC-3’	5’-CTGCCCATCATCATGACCT-3’
Alu	5’-CCTCAATCTCGCTCTCGCTC-3’	5’-CTCTAAGGCTGCTCAATGTCA-3’
McBox	5′-AGGGAATGTCTTCCCATAAAAACT-3′	5′-GTCTACCTTTTATTTGAATTCCCG-3′
Sat a	5′-AAGGTCAATGGCAGAAAAGAA-3′	5′-CAACGAAGGCCACAAGATGTC-3′
Sat III	5′-AATCAACCCGAGTGCAATCNGAATGGAATCG-3′	5′-TCCATTCCATTCCTGTACTCGG-3′

Table 13 . Human RT-qPCR primers

Gene	Forward	Reverse
HOXC13.2	5’-AGCAGAGCTCAGTGGGAGAG-3’	5’-AATTTCAGGCCCACCCTTAG-3’
Globin	5’-GGCTGTCATCACTTAGACCTC-3’	5’-GGTTGCTAGTGAACACAGTTG-3’
Alu	5’-CCTCAATCTCGCTCTCGCTC-3’	5’-CTCTAAGGCTGCTCAATGTCA-3’
McBox	5′-AGGGAATGTCTTCCCATAAAAACT-3′	5′-GTCTACCTTTTATTTGAATTCCCG-3′
Sat a	5′-AAGGTCAATGGCAGAAAAGAA-3′	5′-CAACGAAGGCCACAAGATGTC-3′
Sat III	5′-AATCAACCCGAGTGCAATCNGAATGGAATCG-3′	5′-TCCATTCCATTCCTGTACTCGG-3′

Table 14 . Human ChIP-qPCR primers

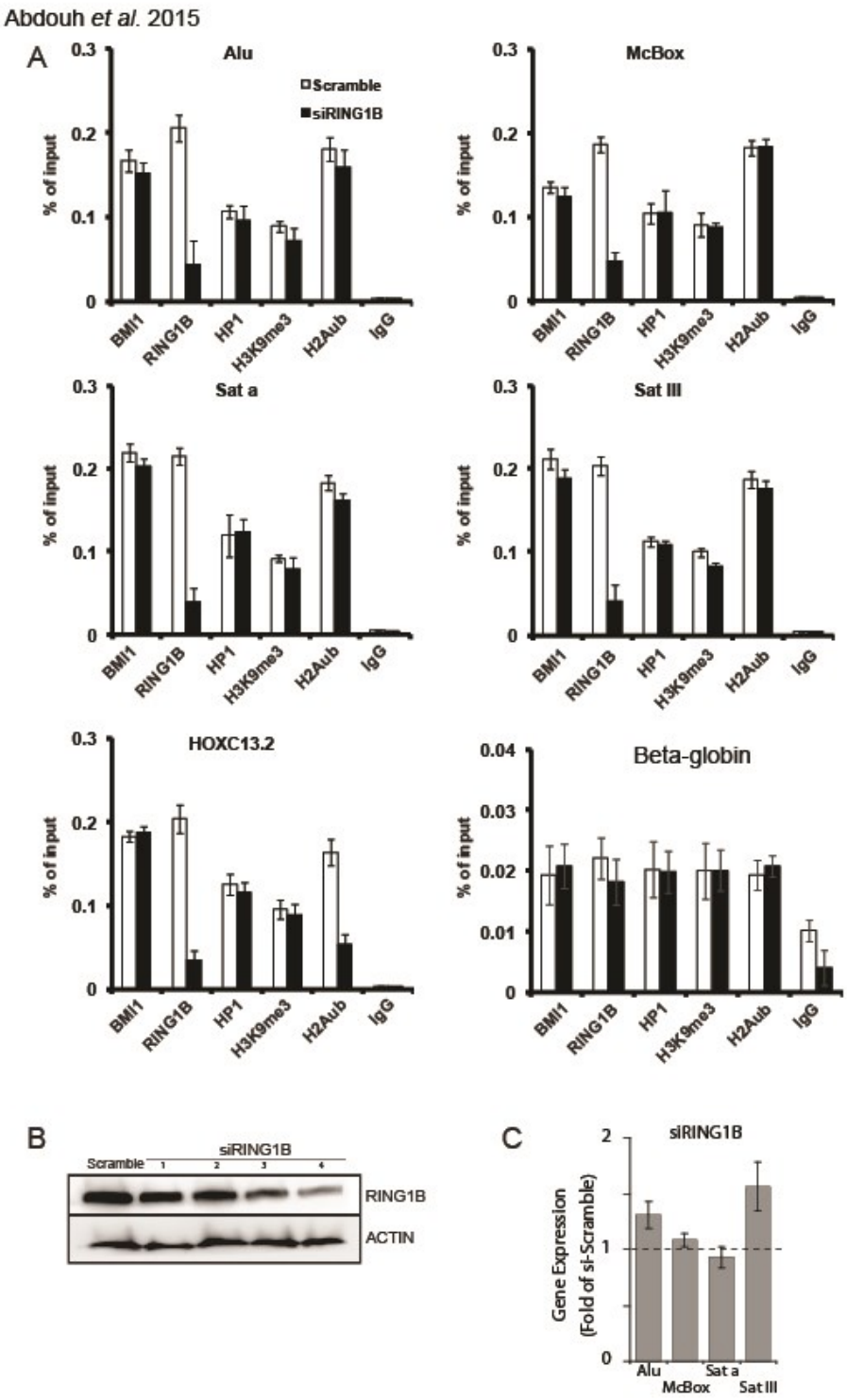


Figure 41. RING1B knockdown does not recapitulate the BMI1-deficient heterochromatin phenotype

(A-C) 293T cells were transfected with siScramble or siRING1B RNAi and analyzed by ChIP, Western blot and qPCR. (A) Although enrichment for RING1B is highly reduced in siRING1B-treated cells at all loci, no significant impact was observed for H2Aub and H3K9me3 deposition and HP1 accumulation at repeat-DNA sequences. Reduced H2Aub deposition was however observed at the canonical target *HOXC13*. (B) The siRING1B # 4 was used for all experiments. (C) RING1B knockdown resulted in modest up-regulation of SatIII expression, although the difference with siScramble was not significant.

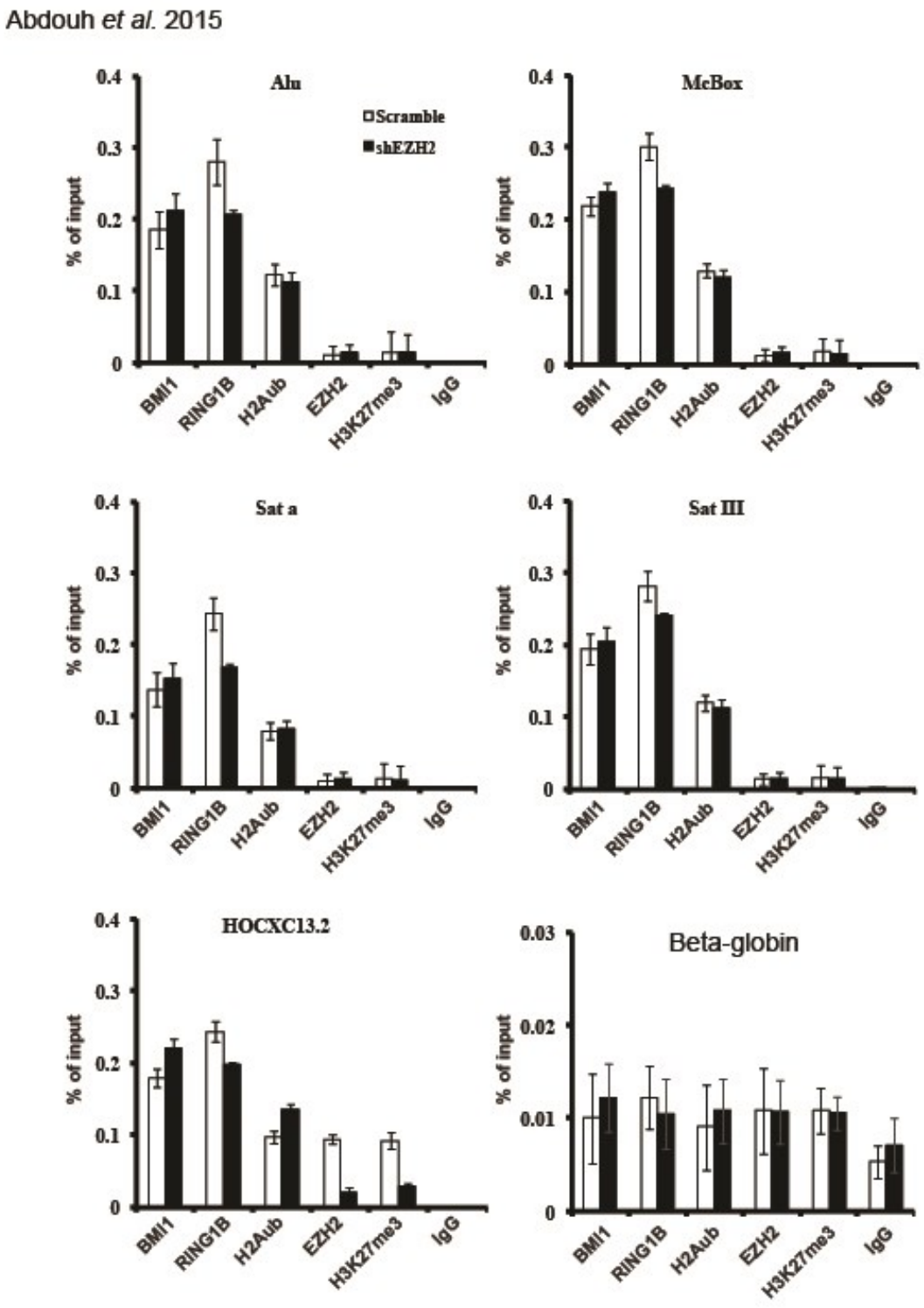


Figure 42. BMI1 localization at repeat-DNA sequences and *HOXC13* is EZH2 and H3K27^{me3}-independent

293T cells were infected with shScramble or shEZH2 viruses and analyzed by ChIP. Although enrichment for EZH2 and H3K27me3 was highly reduced at *HOXC13* in shEZH2-treated cells, there was no impact on BMI1 and H2Aub enrichment. Furthermore, significant enrichment for EZH2 and H3K27me3 was not observed at repeat-DNA sequences in both shScramble- and shEZH2-treated cells.

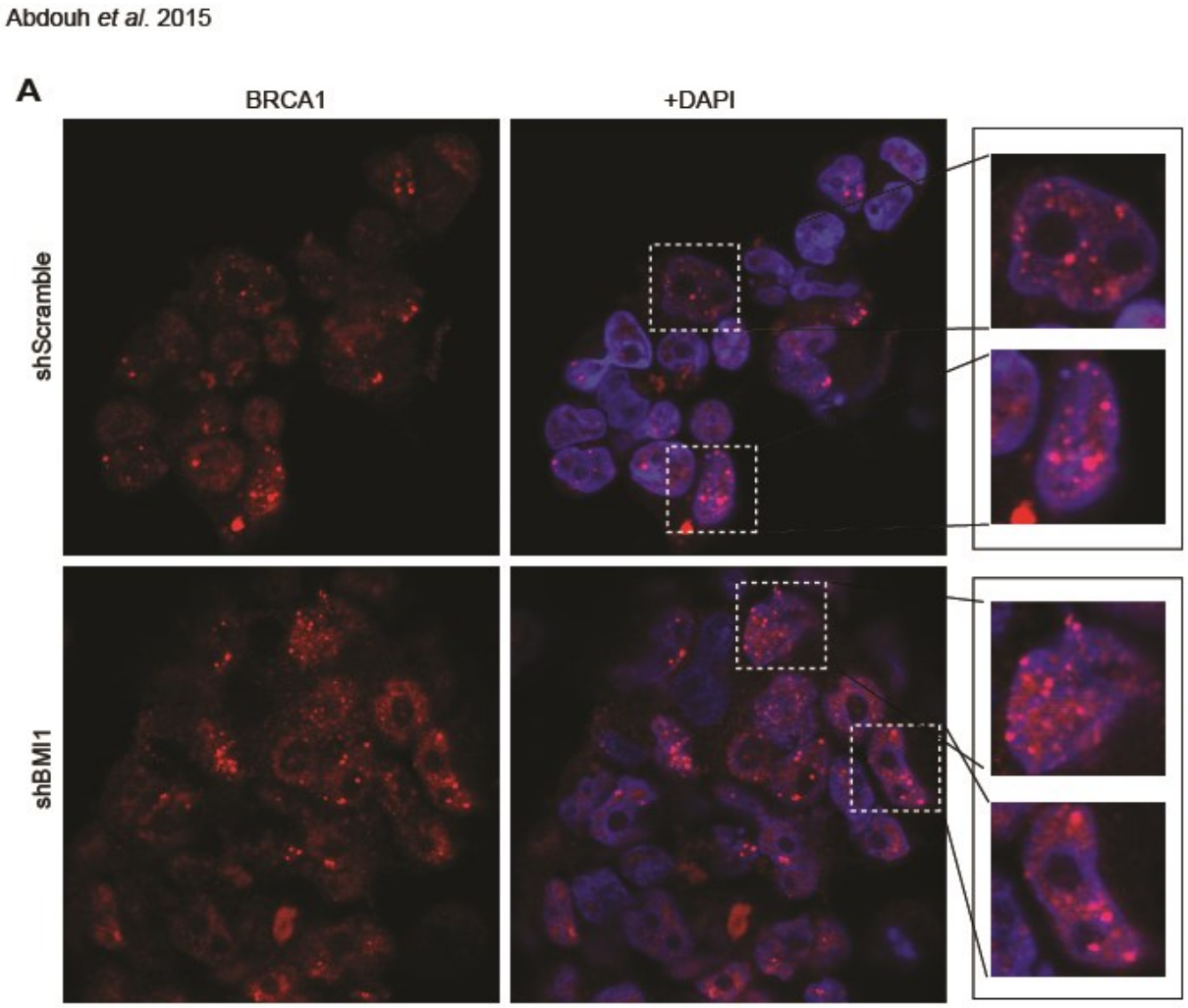


Figure 43. BRCA1 localization is unaffected upon BMI1 deficiency in human cells

293T cells were infected with shScramble or shBMI1 lentiviruses and selected for hygromycin resistance. Immuno-fluorescence analyses confirmed that BRCA1 expression/localization was not affected by BMI1 deficiency.

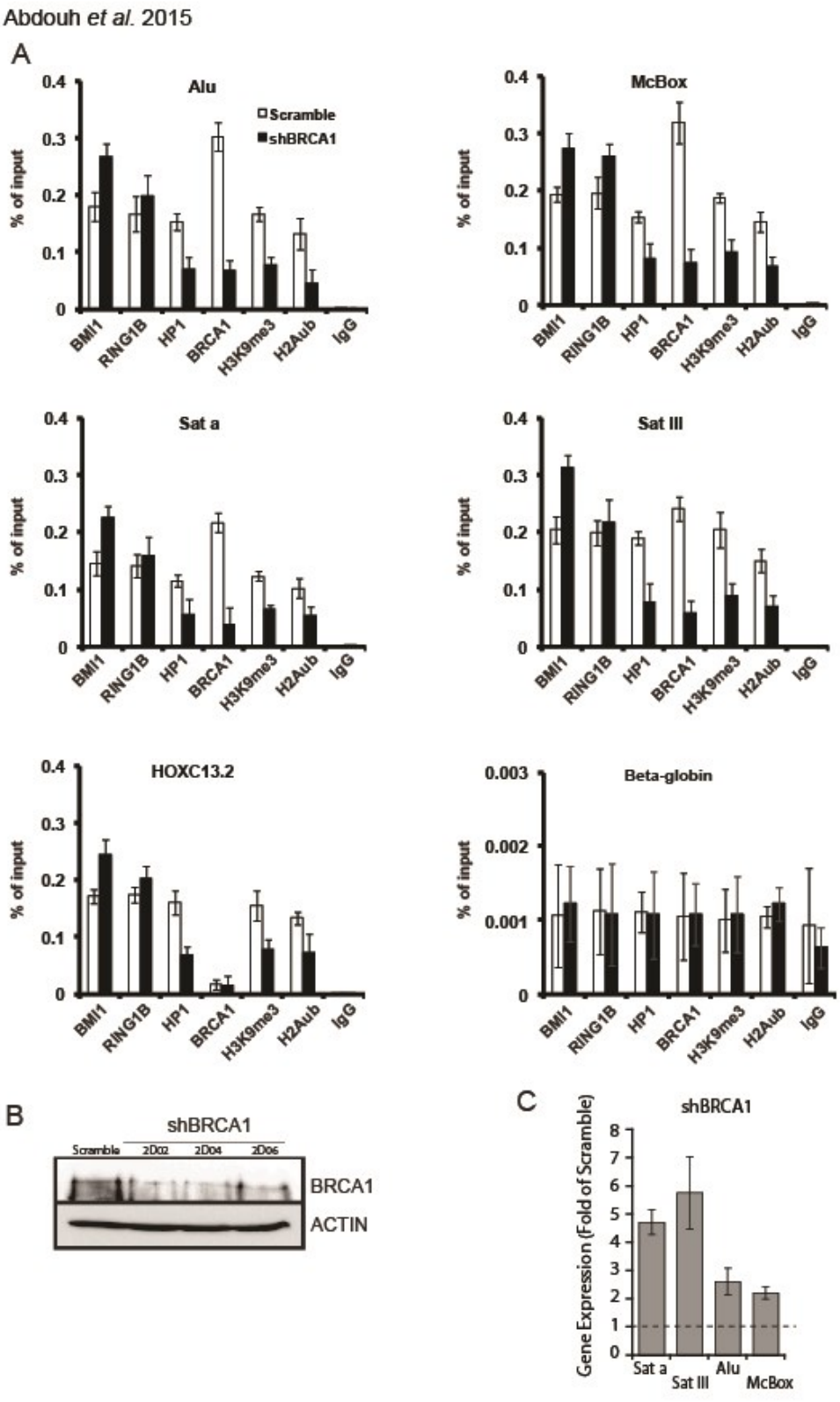


Figure 44. BRCA1 is required for H2Aub deposition, heterochromatin formation and silencing

293T cells were transfected with shScramble or shBRCA1 plasmids and analyzed by ChIP (A), Western blot (B), and qPCR (C). (A) BRCA1 knockdown resulted in reduction of H2Aub and H3K9me3 deposition and HP1 accumulation at all tested loci, including *HOXC13*, where BRCA1 accumulation is relatively low. BRCA1 knockdown also resulted in increased accumulation of BMI1 at all tested loci. (B) The 2D04 shBRCA1 construct was used for all experiments. (C) BRCA1 knockdown resulted in up-regulation of *Satellite repeats*, with a modest effect on *ALU* and *McBox*.

7 Annex II: Heterochromatic genome instability and neurodegeneration sharing similarities with Alzheimer's disease in old Bmi1+/− mice

The first annex is an article, on which I have worked during my Ph.D., and was published on Scientific Reports in 2018. In this article, we have established that BMI1 +/- mice are an interesting animal model to study sAD. These mice, as they grow old, develop many of the molecular, physiological, and behavioral hallmarks of sAD.

SCIENTIFIC REPORTS

OPEN

Heterochromatic genome instability and neurodegeneration sharing similarities with Alzheimer's disease in old *Bmi1*+/− mice

Received: 13 June 2018
Accepted: 30 November 2018
Published online: 24 January 2019

Jida El Hajjar[1], Wassim Chatoo[1], Roy Hanna[1], Patrick Nkanza[1], Nicolas Tétreault[1], Yiu Chung Tse[2,3], Tak Pan Wong[2,3], Mohamed Abdouh[1] & Gilbert Bernier[1,4]

Sporadic Alzheimer's disease (AD) is the most common cause of dementia. However, representative experimental models of AD have remained difficult to produce because of the disease's uncertain origin. The Polycomb group protein BMI1 regulates chromatin compaction and gene silencing. BMI1 expression is abundant in adult brain neurons but down-regulated in AD brains. We show here that mice lacking one allele of *Bmi1* (*Bmi1*+/−) develop normally but present with age cognitive deficits and neurodegeneration sharing similarities with AD. *Bmi1*+/− mice also transgenic for the *amyloid beta precursor protein* died prematurely and present aggravated disease. Loss of heterochromatin and DNA damage response (DDR) at repetitive DNA sequences were predominant in *Bmi1*+/− mouse neurons and inhibition of the DDR mitigated the amyloid and Tau phenotype. Heterochromatin anomalies and DDR at repetitive DNA sequences were also found in AD brains. Aging *Bmi1*+/− mice may thus represent an interesting model to identify and study novel pathogenic mechanisms related to AD.

Alzheimer's disease of the late-onset sporadic form (AD) is the most common dementia[1]. At the genetic level, carriers of the *apolipoprotein E4* (*APOE4*) allele have a higher risk to develop the disease with age[2]. Despite this, the etiology of AD remains largely elusive, with old age representing the greatest risk factor. Familial AD (FAD) is an early-onset, autosomal dominant disorder, that is associated with mutations in *amyloid beta precursor protein* (*APP*), *presenilin 1 (PSEN1) and presenilin 2 (PSEN2)*, and representing less than 5% of all Alzheimer's disease cases[3]. Short-term memory deficits and behavioral changes are the earliest known manifestations of AD and FAD. While brain atrophy is generally restricted to the limbic system and hippocampus at the disease's onset, neurodegeneration eventually spreads to all cortical areas, resulting in loss of up to 30% of the brain mass. At the histo-pathological level, the disease is characterized by the accumulation of amyloid plaques and intra-neuronal phospho-Tau (p-Tau) tangles. Early loss of synapses and, later, of neurons, is also a characteristic of the disease[3]. While numerous transgenic mouse models of FAD have been developed that could recapitulate several hallmarks of the disease, experimental mouse models of AD are rare because the disease's origin remains uncertain[4–6].

Nucleosomes are the basic building unit of chromatin and are constituted of a 147 base pairs of double-stranded DNA wrapped around a histone octamer[7]. Post-translational modification of histones can modify chromatin compaction and stability. BMI1 (B-cell specific Moloney murine leukemia virus integration site 1) is a core component of the Polycomb Repressive Complex 1 (PRC1). The PRC1 can ubiquitylate histone H2A at lysine 119 ($H2A^{ub}$) through the E3-ubiquitin ligase activity of Ring1a/b, its catalytic unit[8–11]. The main classical function of BMI1 is to maintain chromatin compaction and gene silencing at developmental and senescence-associated loci. The *Ink4a* locus, which encodes for the $p16^{Ink4a}$ and $p19^{Arf}$ tumor suppressor proteins, is one of the main targets of BMI1 repressive activity, and cells deficient for *Bmi1* undergo rapid senescence[12,13]. *Bmi1*-null mice present axial skeleton defects, reduced post-natal growth and lifespan, as well as cerebellar degeneration[14]. The BMI1 protein is also recruited at DNA break sites where it can promote DNA damage response (DDR) and repair[15,16]. *Bmi1*-deficiency in mice is furthermore associated with accumulation of DNA damage

[1]Stem Cell and Developmental Biology Laboratory, Maisonneuve-Rosemont Hospital, 5415 Boul. l'Assomption, Montreal, H1T 2M4, Canada. [2]Department of Psychiatry, McGill University, Montreal, Canada. [3]Douglas Mental Health University Institute, Montreal, Canada. [4]Department of Neurosciences, University of Montreal, Montreal, Canada. Correspondence and requests for materials should be addressed to G.B. (email: gbernier.hmr@ssss.gouv.qc.ca)

and reactive oxygen species (ROS), both of which can be improved through co-deletion of *Chk2* or *p53*[17,18]. More recently, BMI1 was found to be abundant at heterochromatin regions containing repetitive DNA sequences (namely the constitutive heterochromatin). *BMI1* knockdown resulted in reduced histone H3 trimethylation at lysine 9 ($H3K9^{me3}$) and in transcriptional activation of tandem repeats in mammalian cells[19].

We recently reported that the expression of *BMI1* is reduced in AD brains and in AD neurons produced from the differentiation of patient-specific induced pluripotent stem cells. Notably, acute *BMI1* inactivation in mature human neurons could recapitulate most AD-related neuronal pathologies[20]. Herein, we show that mice lacking one *Bmi1* allele ($Bmi1^{+/-}$) develop normally, but display with age a pathology sharing some similarities with AD. Crosses between $Bmi1^{+/-}$ mice and transgenic mice carrying a mutant human APP gene revealed genetic cooperation in disease onset and severity. Furthermore, loss of heterochromatin and DDR at repetitive DNA sequences in neurons was characteristic of $Bmi1^{+/-}$ mice and AD brains. Experimental evidences in mice furthermore suggested a significant contribution of the DDR to the observed neuronal pathology. Loss of one *Bmi1* allele in mice thus predisposes to age-related neurodegeneration sharing some similarities with AD.

Methods

Human samples. Paraffin-embedded human brains were obtained from the department of pathology of Maisonneuve-Rosemont Hospital. Frozen post-mortem human cortices are a gift from the Douglas Hospital Brain Bank. The human samples were used accordingly to the Maisonneuve-Rosemont Hospital Ethic Committee (Approval ID #2012-481, 11065). All samples were obtained after informed consent of the patients. Samples were confirmed by histo-pathological analysis as non-demented controls, FAD or AD cases. The list of human samples is described in Supplemental Table 1.

Animals. Mice (male and female) were used in compliance with- and with the approval of the Animal Care Committee of the Maisonneuve-Rosemont Hospital Research Center (Approval ID #2009-40; #2009-42; # 2011-23). The FAD mouse model (Tg-Thy-1APPSwDutIowa in C57BL/6 J background) was purchased from The Jackson laboratory[21]. $Chk2^{+/-}$ mice were obtained from N. Motoyama. $p53^{+/-}$ and $Ink4a^{+/-}$ mice were obtained from The Jackson Laboratory[17,22]. $Bmi1^{+/-}$ mice were a generous gift from M. van Lohuizen (Netherlands)[14]. All mice were maintained in the C57BL/6 J genetic background. $Bmi1^{+/-}$ mice were crossed with either $Chk2^{+/-}$ or $p53^{+/-}$ mice, to obtain $Bmi1^{+/-}/Chk2^{+/-}$ or $Bmi1^{+/-}/p53^{+/-}$ mice. $Bmi1^{+/-}/Chk2^{+/-}$ or $Bmi1^{+/-}/p53^{+/-}$ mice were then intercrossed to obtain $Bmi1^{-/-}/Chk2^{-/-}$ or $Bmi1^{-/-}/p53^{-/-}$ mice. From several breeding, we obtained the expected Mendelian ratio for $Bmi1^{-/-}/Chk2^{-/-}$ mice (12.5%), but not for $Bmi1^{-/-}/p53^{-/-}$ mice (1 animal out of 200 offspring or 0.5%), suggesting that the *Bmi1/p53* double mutant is lethal[23]. A similar strategy was used to generate $Bmi1^{-/-}/Ink4a^{-/-}$ mouse embryos in order to perform the neuronal cultures. Wild type ($Bmi1^{+/+}$) and $Bmi1^{-/-}$ animals used for comparison were littermates of the double mutants.

Murine neuronal cultures. After dissociation of E18.5 cortices, cells were plated at 1.5×10^5 cells/well on poly-L-lysine-coated 8-well cultures slides (BD Biosciences). Cells were maintained in normal medium composed of Neurobasal-A Medium (Invitrogen), Glutamax-I (Gibco), gentamycin (50 µg/ml; Gibco), B27 supplement (Gibco), nerve growth factor (50 ng/ml; Invitrogen) and BDNF (0,5 ng/ml; Invitrogen). For the assays using inhibitors, we added 2 µM of DMSO, 2 µM of ATM/ATR inhibitor (CGK733; Millipore) or 10 µM of ATM inhibitor (KU55933; KuDOS Pharmaceuticals) to the cell culture medium for 16 hours. NAC was added daily at 5 mM into the culture medium for 7 days[18].

Amyloid fractions and Immunoprecipitation. Cortices were minced and homogenized in a 10 ml TBS 1 × (50 mM Tris-Cl, pH 7.5. 150 mM NaCl)/protease inhibitor cocktail solution (Roche Diagnostics), using a tissue grinder. The homogenates were sonicated and ultra-centrifuged at 175,000 g for 1 h at 4 °C. The supernatants were collected as the soluble cellular fraction[24]. The dense pellet was re-homogenized in a 10 ml TBS 1×/1% Triton X-100/protease inhibitor cocktail solution, followed by sonication and centrifugation at 100,000 g for 1 h at 4 °C. The supernatants were recuperated as the insoluble cellular fraction[24]. For immunoprecipitation, protein samples and cellular fractions were incubated overnight with the primary antibody followed by incubation with the protein A/G beads (Millipore) for 2 hours the second day. Beads with bound immunocomplexes were washed with RIPA buffer (50 mM Tris-HCl, pH 7.5; 150 mM NaCl; 0.1%Tween 20; protease inhibitors Complete) and bound proteins were subsequently heat eluted with 1x Laemmli buffer.

Western blot. Cortices were minced and homogenized in 5 ml of complete mini protease inhibitor cocktail solution (100 mM Tris-HCl, pH7.5; 150 mM NaCl; 0.1% Tween 20; 2% SDS; protease inhibitors Complete; Roche Diagnostics) using a tissue grinder. The homogenates were sonicated and quantified using the Bradford reagent. For Western Blot, proteins were resolved by SDS-PAGE electrophoresis and transferred to a nitrocellulose blotting membrane (Bio-Rad). Membranes were incubated with the primary antibodies, treated with the corresponding horseradish peroxidase-conjugated secondary antibodies (Sigma), and developed using the Immobilon Western (Millipore).

Immunohistochemistry. Tissues were immersed in Formalin or 4% paraformaldehyde/3% sucrose in 0.1 M phosphate buffer, pH 7.4 for 16 h at room temperature on a rocking platform. Samples were washed three times in PBS and embedded in paraffin according to standard protocols. 8 µm thick sections were mounted on Super-Frost glass slides (Fisher Scientific) and processed for immunohistochemistry staining. Paraffin-embedded slices were analyzed by using the Vectastain® ABC kit (Vector) according to the manufacturer instructions. Peroxidase substrates used are the Vector® VIP (Violet) (Vector), and DAB (brown) (Sigma)[18]. Observations were made using the Zeiss imager Z2 microscope and images were captured with a digital camera. To obtain comparable results, we selected pyramidal neurons from the frontal cortex to quantify the number of chromocenter/neuron. Pyramidal

neuron identity was determined by the expression of NeuN, the cell morphology (using phase contrast imaging), the large nuclear diameter (when compared to astrocytes), and the presence of chromocenters and of a large nucleolus.

Senescence-associated β-galactosidase assay. Senescence-associated (SA) β-galactosidase staining was detected histochemically at pH 6 as described previously[25]. Briefly, tissue slices were fixed in PBS containing 1% formaldehyde, 0.4% glutaraldehyde, and 0.02% Igepal. After three washes, slices were exposed to the X-gal solution (1 mg/ml X-gal; 5 mM K3Fe(CN)6; 5 mM K4Fe(CN)6; 1 mM MgCl2, in PBS; pH 6.0). Slides were mounted and observations made under a microscope.

Antibodies. Rabbit anti-p53 (Santa Cruz Biotechnology), rabbit anti-p-JNK (Invitrogen), mouse anti-synaptophysin (Sigma), mouse anti-amyloid clone DE2B4 (Abcam), rabbit anti-amyloid clone FCA3542 (Calbiochem), mouse anti-amyloid MOAB2 (Novus; NBP2-13075)), pan anti-Tau k9JA (DAKO); mouse anti-p-Tau clone AT-8 (Thermo scientific), mouse anti-p-Tau clone PHF1 (a gift from Dr. Davies, Albert Einstein College of Medicine), mouse anti-NeuN (Chemicon), rabbit anti-cleaved caspase-3 (Cell Signaling), rabbit anti-GFAP (Dako; Z0334), Iba1 (Wako; 019-19741), mouse anti-Bmi1 clone F6 (Millipore), mouse anti-Bmi1 (Abcam), mouse anti-p-ATM (Novus), rabbit anti-p-ATR (Santa Cruz Biotechnology), rabbit anti-H3K9^{me3} (Abcam), mouse anti-H2Aub clone E6C5 (Millipore), mouse anti-HP1 (Millipore), mouse anti-β-actin (Sigma), mouse anti-tubulin (Sigma), mouse anti-H3 (Upstate), and rabbit anti-mouse IgG (Upstate).

RT and real-time PCR. All primers were designed to flank individual exons and tested by PCR in RT+ and RT− control extracts. Total RNA was isolated using TRIzol reagent (Invitrogen). The RNA was treated with DNaseI for 20 min prior to reverse transcription (RT) and PCR for satellite repeats, which are intron-less. Reverse transcription was performed using 1 µg of total RNA and the MML-V reverse transcriptase (Invitrogen). Real-time PCR was performed using the Platinum SYBRGreen SuperMix (Invitrogen) and a Real-Time PCR apparatus (BioRad). Primers sets used were as follow: p16Ink4a (F) 5_-CAACGCCCCGAACTCTTTC-3_; (R) 5_-GCAGAAGAGCTGCTACGTGAAC-3_, p19Arf (F) 5_-GGCTAGAGAGGATCTTGAGAAGAGG-3_; (R) 5_-GCCCATCATCATCACCTGGTCCAGG-3_. Primers sets used for Bmi1, IAP, Maj Sat, Min Sat, Line, Sine and IAP were as described[19].

Chromatin Immunoprecipitaion (ChIP) assay. ChIP was performed using the ChIP Assay kit (Upstate). Briefly, 50 mg of cortical tissue was frozen for 1 hour and then homogenized at RT according to the manufacturer's protocol. The tissue was sonicated on ice for 10 sec at 30% amplitude to shear the chromatin (Branson Digital Sonifier 450, Crystal Electronics, On. Canada). Sonicated materials were immunoprecipitated using designated antibodies. Fragments were then amplified by real-time PCR in triplicates. Human primers sets used were as follow: MCBOX (F) 5′-AGGGAATGTCTTCCCATAAAAACT-3′; (R) 5′-GTCTACCTTTTATTTGAATTCCCG-3′; SATIII (F) 5′-AATCAACCCGAGTGCAATCNGAATGG-3′; (R) 5′-TCCATTCCATTCCTGTACTCGG-3′; SATa (F) 5′-AAGGTCAATGGCAGAAAAGAA-3′; (R) 5′-CAACGAAGGCCACAAGATGTC-3′; ACTIN (F) 5′-CCTCAATCTCGCTCTCGCTC-3′; (R) 5′-CTCTAAGGCTGCTCAATGTCA-3′; ß-GLOBIN (F) 5′-GGCTGTCATCACTTAGACCTC-3′; (R) 5′-GGTTGCTAGTGAACACAGTTG-3′;5′. Mouse primers sets used were as follow: MajSAT 5′- GGCGAGAAAACTGAAAATCACG-3′, 5′-CTTGCCATATTCCACGTCCT-3′; MinSAT 5′- TTGGAAACGGGATTTGTAGA-3′, 5′-CGGTTTCCAACATATGTGTTTT-3′; LINE 5′-TGGCTTGTGCTGTAAGATCG-3′, 5′-TCTGTTGGTGGTCTTTTTGTC-3′; SINE1 5′-GAGCACACCCATGCACATAC-3′, 5′-AAAGGCATGCACCTCTACCACC-3′; Actin 5′-TCGATATCCACGTGACATCCA-3′; 5′-GCAGCATTTTTTTACCCCCTC-3′; HoxA7 5′-GTGGGCAAAGAGTGGATTTC-3′; 5′-CCCCGACAACCTCATACCTA-3′; β-major 5′-CAGTGAGTGGCACAGCATCC-3′; 5′-CAGTCAGGTGCACCATGATGT-3′. ChIP-qPCR data was analyzed according to the Percent Input method[19]. First, the raw Ct of the diluted 1% input fraction is adjusted by subtracting 6.64 cycles (i.e. log2 of the dilution factor 100). Subsequently, the percent input of each IP fraction is calculated according to this equation: $100*2^{(\text{Adjusted Input Ct-Ct(IP)})}$.

Behavioral tests. Behavioral assays were performed at the Douglas Hospital Institution, Montreal, Canada under the supervision of Dr. Rocheford. Male mice were first accustomed to the test environment prior to testing. Groups of 15-month old male WT (n = 8) and *Bmi1*$^{+/-}$ (n = 8) mice were tested successively and were habituated, trained, and evaluated at the same time of day for each experiment. Behavioral tests included Morris water maze test and Barnes maze assay. In the Morris water maze assay, a hidden escape platform is immersed just below water level with visual cues (color) around the pool in plain eyesight of the mouse. Upon release, the mouse navigates around the pool seeking an exit while various parameters are recorded, including the time spent in each quadrant of the pool and the time taken to reach the platform (latency). Barnes maze assay includes a circular platform containing 20 circular holes around its circumference and visual cues (colored shapes) surrounding the surface in plain eyesight of the mouse. The platform is intensely lit by an overhead lamp, and underneath one of the holes there is an exit box. Different parameters are analyzed comprising latency to escape and number of errors.

LTP assay. For this study, we used 15-month old male WT and *Bmi1*$^{+/-}$ mice. Under deep anesthesia, brains were removed and coronal brain slices (350 µm thickness) containing the hippocampus were cut in hyperosmotic, ice-cold and carbogenated (bubbled by 95%O_2/5%CO_2 to maintain the pH at 7.4) slice cutting solution (in mM: 252 sucrose, 2.5 KCl, 4 $MgCl_2$, 0.1 $CaCl_2$, 1.25 KH_2PO_4, 26 $NaHCO_3$ and 10 glucose) using a Vibratome. Freshly cut slices were placed in an incubating chamber with carbogenated artificial cerebrospinal fluid (ACSF (~310 mOsmol/L) consists of (mM): 125 NaCl, 2.5 KCl, 1 $MgCl_2$, 2 $CaCl_2$, 1.25 NaH_2PO_4, 26 $NaHCO_3$ and 25 glucose).

Slices were recovered at 32 °C for one hour and subsequently maintained at room temperature. Carbogenated ACSF containing bicuculline methbromide (5 µM) to block $GABA_A$ receptor-mediated inhibitory synaptic current was used to perfuse slices in all recordings. Postsynaptic responses, which were evoked by stimulating the Schaffer collateral-commissural pathway via constant current pulses (0.08 ms) delivered through a tungsten bipolar electrode (FHC), were recorded from the hippocampal CA1 region, amplified by a Multiclamp 700B (Axon) and stored in a PC for offline analysis using Clampfit (Axon). All recordings were performed at room temperature. Field excitatory postsynaptic potential (fEPSP) was evoked at 0.05 Hz and detected by an ACSF-filled glass electrode that was placed in the stratum radiatum of the hippocampal CA1 region. Long-term potentiation (LTP) of fEPSP was induced by a high-frequency tetanus (100 Hz, 100 pulses). Student's t-test was used to compare changes in fEPSP slope at 60 minutes after tetanus between WT and *Bmi1*$^{+/-}$ groups. In order to determine if fEPSPs are potentiated by tetanus within each mouse group, paired Student's t-test was used to compare fEPSP slope recorded before (baseline) and at 60 minutes after tetanus.

Statistical analysis. The Student's *t*-test for unpaired samples was used to measure statistical differences. An analysis of variance (ANOVA) followed by the Dunnett test was used for multiple comparisons with one control group. In each figures, the criterion for significance (*P* value) was mentioned.

Results

***Bmi1*$^{+/-}$ mice develop accelerated aging features and neurological symptoms.** We previously reported that *Bmi1*$^{+/-}$ mice have reduced median and maximal lifespan along with apparition of lens cataracts at higher frequencies than WT littermates, suggesting that *Bmi1* hemi-deficiency may accelerate the aging process[18]. Alopecia and weight loss are common characteristics of premature aging in mice[26–30]. We observed that between the ages of 15–24 months (defined here as *old* mice), *Bmi1*$^{+/-}$ mice exhibited alopecia and reduced body size (~25%) when compared with age-match WT littermates (Fig. 1a,b). These pathological features were not present in 3-month and 1-year old *Bmi1*$^{+/-}$ mice (Fig. 1a,b). We performed the paw-clasping test to verify if old *Bmi1*$^{+/-}$ mice presented pathological reflexes. Adult WT mice picked by the tail and slowly inclined towards a horizontal surface spread all four limbs in anticipation of contact. However, instead of limb extension, old *Bmi1*$^{+/-}$ mice displayed an irregular clasping behavior characterized by the dragging of the fore paws towards the trunk and asymmetry in limb position (Fig. 1c). This phenotype is generally associated with a neurological syndrome[31].

The Bmi1 gene operates as a central regulator of apoptosis and senescence through repression of the $p19^{Arf}$/p53 and $p16^{Ink4a}$/pRb tumor suppressor axes. Senescence is usually excluded in post-mitotic cells such as neurons, yet increased SA β-galactosidase activity and expression of the senescence-associated genes $p19^{Arf}$ and $p16^{Ink4}$ have been noted in cortical neurons of aged *p73*$^{+/-}$ mice and of *Bmi1*$^{-/-}$ pups[18,25,32]. Furthermore, expression of $p16^{Ink4}$ is considered as a biomarker of aging[33]. We first confirmed neuronal expression of Bmi1 in the cortex of old WT mice and its severe reduction in old *Bmi1*$^{+/-}$ mice (Fig. 1d). Next, we compared these old WT and *Bmi1*$^{+/-}$ mice for SA β-galactosidase activity and $p16^{Ink4}$/$p19^{Arf}$ expression. We observed robust SA β-galactosidase staining in pyramidal neurons and a 2-fold increase in the total number of SA β-galactosidase-positive neurons/section in layers 2 and 3 of the frontal cortex of *Bmi1*$^{+/-}$ mice (Fig. 1e,f). This was accompanied with significant increased in $p16^{Ink4}$ and $p19^{Arf}$ expression (Fig. 1g). These findings suggested that in addition to affect lifespan, *Bmi1* hemi-deficiency resulted in the apparition of an age-dependent neurological syndrome associated with neuronal senescence.

***Bmi1*$^{+/-}$ mice show altered spatial memory and reduced long-term potentiation.** To test whether the neurological symptoms were associated with perturbed cognitive functions, we examined apparently healthy 15 month-old WT and *Bmi1*$^{+/-}$ mice for spatial learning and memory aptitude using the Morris water and Barnes maze tests. For the Morris water navigation task, mice were trained for 4 days to locate the platform (training) and then evaluated on their capacity to remember the platform location (probe test). *Bmi1*$^{+/-}$ mice performed poorly in this assay during both the training period and probe test, as demonstrated by the increased time (latency) required to reach the platform (Fig. 2a,b). *Bmi1*$^{+/-}$ mice also exhibited an anxious phenotype as demonstrated by an increased thigmotactic behavior while performing the training (Fig. 2d)[34]. The swimming speed of aged *Bmi1*$^{+/-}$ mice was also reduced (0.25 m/s for WT and 0.20 m/s for *Bmi1*$^{+/-}$) (Fig. 2c). In contrast, the swimming speed of 1 year-old WT and *Bmi1*$^{+/-}$ mice was nearly identical (not shown), thus consistent with the apparition of a progressive, age-related reduction in locomotors strength. The Barnes maze test is based on mice aversion to open enlightened spaces, which ultimately prompts them to find secure shelter. *Bmi1*$^{+/-}$ mice were more prone to visiting non-target holes while searching for the hole with the escape box and the latency of escape was also increased (Fig. 2e). Although these results suggested impaired spatial memory formation, the reduction in swimming speed and the possibility of visual dysfunction[23] led us to perform electrophysiological analyses to further test this possibility.

LTP is a cellular surrogate for spatial memory formation[35]. Using the same animals as in the behavioral tests, LTP of the Schaffer collateral pathway was induced in the CA1 region of the hippocampus. We found that a commonly used protocol for inducing LTP (100 Hz, 100 pulses) enhanced fEPSP slope in WT mice at 60 minutes after tetanus (Percent potentiation: 48.9 ± 5.2%, data from 6 slices and 3 mice). Compared with fEPSP recorded before tetanus (baseline), WT slices displayed significant potentiation ($P = 0.036$; paired Student's t-test) (Fig. 2f–h). Slices from *Bmi1*$^{+/-}$ mice exhibited weaker potentiation than WT (11.7 ± 11.2%, data from 6 slices, 4 mice; transgenic vs. WT: $P = 0.045$, Student's t-test) and failed to show significant potentiation when compared to baseline at 60 minutes after tetanus ($P = 0.357$; paired Student's t-test) (Fig. 2f–h). Our findings demonstrated that LTP formation in *Bmi1* hemi-deficient mice was significantly impaired, thus strengthening the hypothesis of impaired spatial memory formation.

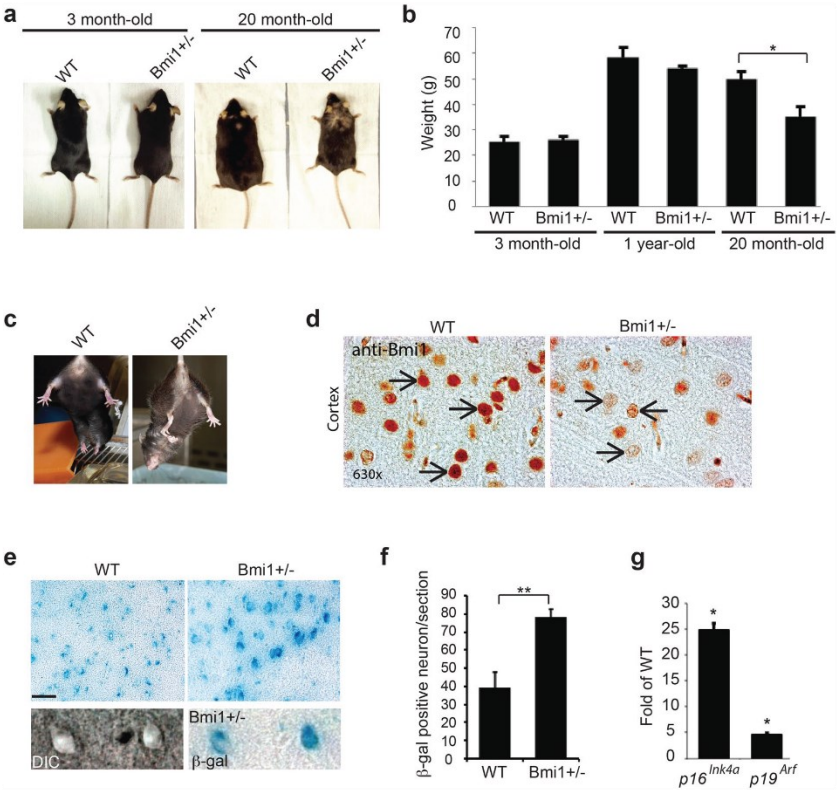


Figure 1. *Bmi1*$^{+/-}$ mice present accelerated aging features and neurological symptoms. (**a**) Photographs of 3 month-old and 20 month-old WT and *Bmi1*$^{+/-}$ mice. Note that aged *Bmi1*$^{+/-}$ mice are smaller and exhibit hair loss (alopecia) compared to WT littermates (n = 5). (**b**) Quantification representing the mean weight at 3 months (n = 6), 1-yr (n = 6) and 20 months (n = 6) of WT and *Bmi1*$^{+/-}$ mice. (**c**) Photographs showing the paw clasping reflex of aged WT (n = 5) and *Bmi1*$^{+/-}$ mice (n = 5). Note that the *Bmi1*$^{+/-}$ mouse pulls its fore limbs into its body rather than spreading them away from its trunk, and present an important asymmetry in the hind limbs' position. (**d**) IHC analysis on mouse cortical sections showing robust Bmi1 expression in the nucleus of WT neurons (arrows) and weak expression in *Bmi1*$^{+/-}$ neurons (arrows). (**e**) Analysis of SA β-galactosidase activity in the cortices of old WT and *Bmi1*$^{+/-}$ mice. Scale bar: 20μm. (**f**) Quantification of the number of SA β-galactosidase positive neurons/cortical section. (**h**) Quantitative PCR analysis of *p16*^{Ink4a} and p19^{Arf} expression in the cortices of old WT (n = 3) and *Bmi1*$^{+/-}$ (n = 3) mice.

Old *Bmi1*$^{+/-}$ mice present a pathology showing some features of AD. We previously showed that most neuronal anomalies in *Bmi1*$^{-/-}$ mice were mediated by p53 activation[18]. Using immuno-blot, we observed increased p53 levels in the cortex of 15–20 month-old *Bmi1*$^{+/-}$ mice when compared to WT littermates (Fig. 3a). Hallmarks of AD and related dementias are accumulation of p-Tau, extra-cellular amyloid plaques, reduced synaptophysin immunoreactivity and neuronal loss. By immunoblot using antibodies that recognize p-Tau at Ser202 (AT-8) or Ser396 and Ser404 (PHF1), we observed the presence of p-Tau accumulation in *Bmi1*$^{+/-}$ cortices (Fig. 3a). In cultured human neurons, *BMI1* knockdown results in an increase in total Tau levels[20]. Similarly, we observed that total Tau levels were increased in the cortex of 22–24 month-old *Bmi1*$^{+/-}$ mice (Fig. S1a,b). Remarkably, these anomalies correlated with reduced synaptophysin levels (indicative of synaptic atrophy) and increased p-JNK immunoreactivity (Fig. 3a)[36,37]. By immuno-histochemistry (IHC), we observed p-Tau immunoreactivity in the neuron's perikarya of 20 month-old WT and *Bmi1*$^{+/-}$ mice. However, large p-Tau deposits in the neuronal cell body and strong immunoreactivity in track fibers of the cortical white matter were only observed in *Bmi1*$^{+/-}$ mice (Fig. 3b). The presence of p-Tau tangles was not observed.

We next performed immunoblot on mouse cortices using an antibody (FCA3542) that can recognize the mouse and human amyloid peptide. Notably, we could detect the presence of a ~12–14 kDa band only in the *Bmi1*$^{+/-}$ brain samples (Fig. 3c-Input 20%). This band was also revealed using FCA3542 after

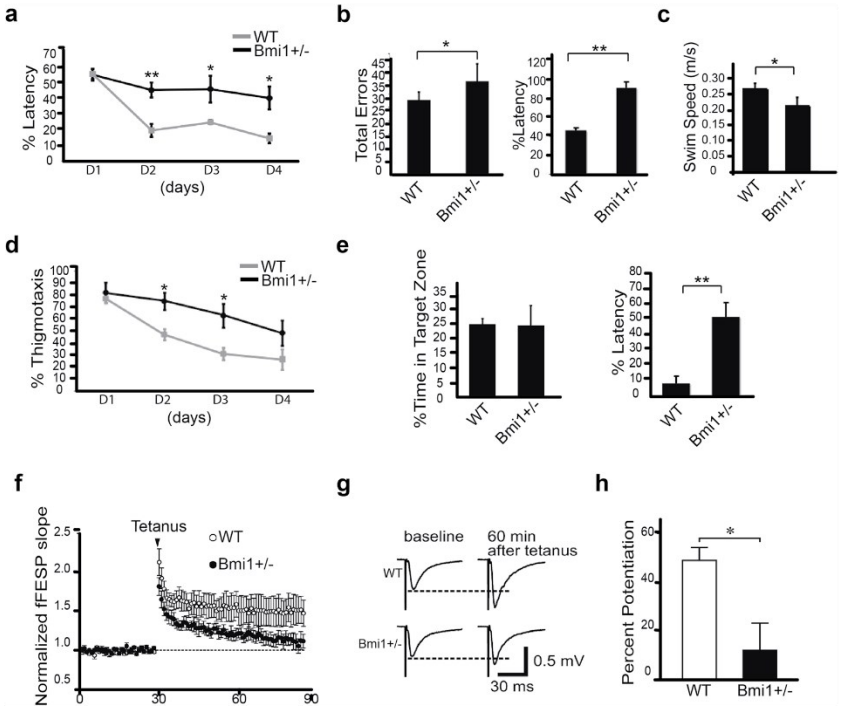


Figure 2. *Bmi1*$^{+/-}$ mice exhibit memory and LPT impairment. (**a–e**) 15-month old male WT (n = 8) and *Bmi1*$^{+/-}$ (n = 8) mice were tested for behavioral changes associated with spatial memory formation deficit. (**a–c**) Results of Morris water maze probe test, showing time spent in the target quadrant searching for the hidden platform, mean time to reach the target platform (latency) and mean swim speed. (**d**) Charts representing mean thigmotaxis of Morris water maze 4-day trials. (**e**) Histograms representative of Barnes maze test results including mean total errors and mean latency to reach the target probe. (**f–h**) LTP assays on hippocampal slices of 15-month old male WT and *Bmi1*$^{+/-}$ mice. (**f**) Scatter plots revealed changes in the slope of field excitatory postsynaptic potential (fEPSP) recorded from the hippocampal CA1 region of WT and *Bmi1*$^{+/-}$ mice. LTP was induced by tetanus (100 Hz, 100 stimulation). Dotted line represents the baseline. (**g**) Representative traces of fEPSP recorded before (baseline) and 60 min after tetanus. Dotted lines represent baselines. (**h**) Histogram of fEPSP slope at 60 min after LTP induction. Normalized fEPSP slopes, recorded between 55–60 min after tetanus, were averaged for calculating percent potentiation. Values are mean ± SEM. *P < 0.05; **P < 0.01; Student's t-test.

immunoprecipitation (IP) with another anti-amyloid antibody (DE2B4), but not after IP with IgG (Fig. 3c-IP). Antibody specificity was also confirmed using cortical extracts from human *APP*$^{Swe/Dutch/Iowa}$ (*APP*) transgenic mice (Fig. 3c-IP). On *Bmi1*$^{+/-}$ brain sections at 20 months of age, we observed that amyloid was preferentially found in the neuronal soma (Fig. 3d). Amyloid plaques were not observed. Cleavage of APP by the β-secretase Bace1 generates the ~12–14 kDa C99 fragment (C99). Notably, we observed increased Bace1 levels in cortical neurons of *Bmi1*$^{+/-}$ mice (Fig. 3a,e). To distinguish between accelerated aging and neurodegeneration, we compared the brains of 24 month-old WT mice with that of 15 month-old *Bmi1*$^{+/-}$ mice for the expression of the ~12–14 kDa fragment. Notably, the amyloid-related band was not detectable in the cortex of old WT mice (Fig. 3f). Likewise, p-JNK and p-Tau (AT-8) were not observed in the 24 month-old WT mice (Fig. 3g).

To evaluate the possibility of neuronal loss, we calculated the total number of cortical neurons per section in the frontal cortex using a pan-neuronal marker (NeuN). When compared to WT littermates, 15–20 month-old *Bmi1*$^{+/-}$ mice displayed a ~20% reduction in NeuN-positive neurons (Fig. 3h). To examine if the lower neuronal density observed in *Bmi1*$^{+/-}$ mice correlated with increased apoptosis, we quantified the number of neurons positive for activated caspase-3 in the frontal cortex. This revealed a ~2 fold increase in the frequency of apoptotic neurons in old *Bmi1*$^{+/-}$ mice (Fig. 3i). Reduced neuronal density and increased apoptosis were also observed in the hippocampus of 15–17 month-old *Bmi1*$^{+/-}$ mice (Fig. S2a,b), which was consistent with the behavioral and LTP results. We next analyzed the brains from rare surviving 22–25 month-old *Bmi1*$^{+/-}$ mice using antibodies against amyloid (n = 7). A single 24 month-old *Bmi1*$^{+/-}$ animal presented amyloid deposits that were present in all cortical areas, with other brain regions being generally less affected (Fig. 3j). Interestingly, the

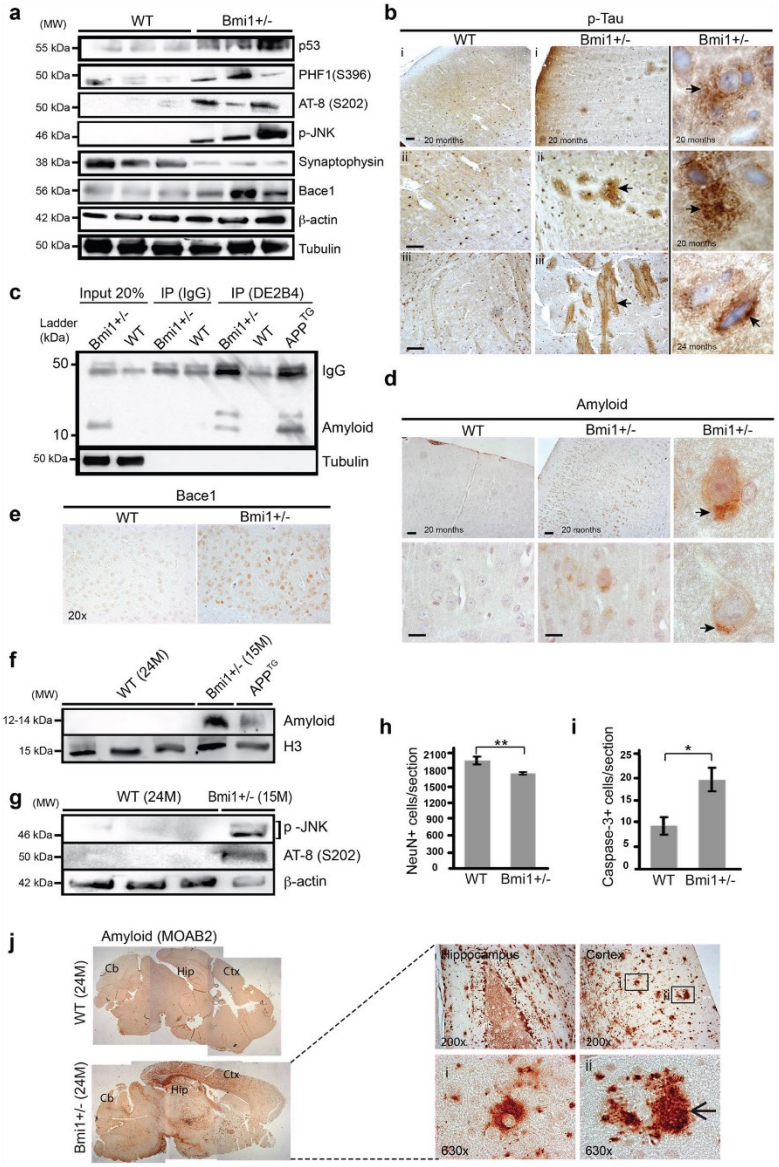


Figure 3. Old *Bmi1*$^{+/-}$ mice present cortical neurodegeneration. (**a**) Western blot analysis on cortical extracts from 15–20 month-old WT and *Bmi1*$^{+/-}$ mice. (**b**) IHC for p-Tau (PHF1) on cortical sections of old WT (n = 4) and *Bmi1*$^{+/-}$ (n = 4) mice. (i) Low magnification. Scale bar: 20μm. Higher magnifications in (ii) revealed p-Tau deposits on ghost-like neurons and in (iii) p-Tau immunoreactivity in track fibers of the cortical white matter of *Bmi1*$^{+/-}$ mice. The 3 insets show images of the most severe cases observed of p-Tau accumulation in *Bmi1*$^{+/-}$ mouse neurons (arrows). Scale bar: 8μm. (**c**) Immuno-precipitation (IP) of the β-amyloid peptide (DE2B4 antibody) using the soluble cellular fraction of cortices from old WT and *Bmi1*$^{+/-}$ mice, and from a 6-month old *APP* transgenic mouse (positive control). The blot was revealed using the FCA3542 antibody. N = 2 independent experiments. (**d**) IHC on brain sections of old WT (n = 4) and *Bmi1*$^{+/-}$ (n = 4) mice for the expression of β-amyloid (DE2B4 antibody). Top images scale bar is 20μm; Lower images scale bar is 8μm. Arrows in the insets indicate amyloid accumulation in the neuronal cell body. (**e**) IHC staining for Bace1 on brain sections of old WT (n = 2) and *Bmi1*$^{+/-}$ (n = 2) mice. (**f**,**g**) Immuno-blot analysis of cortical extracts from 24 month-old WT and 15 month-old *Bmi1*$^{+/-}$ mice showing accumulation of pathological proteins in

Bmi1$^{+/-}$ mice but not in old WT mice. N = 2 independent experiments. (**h**) Quantification of the number of NeuN + cortical neurons in the frontal cortex of 15–20 month-old WT (n = 5) and *Bmi1*$^{+/-}$ (n = 5) using IHC. (**i**) Quantification of the number of neurons positive for activated caspase-3 in the frontal cortex of 15–20 month-old WT (n = 5) and *Bmi1*$^{+/-}$ (n = 5) mice. (**j**) Composite images of IHC staining on whole brain sections from 24-month old WT and *Bmi1*$^{+/-}$ mice showing amyloid-positive deposits in the cortex (Ctx) and hippocampus (Hip) but not in the cerebellum (Cb) of *Bmi1*$^{+/-}$ mice. (Right panel) High-resolution images showing extra-cellular amyloid plaque-like structures (arrow) in the cortex and hippocampus of *Bmi1*$^{+/-}$ mice. Values are mean ± SEM. *P < 0.05; **P < 0.01; Student's t-test.

pattern of amyloid immuno-reactivity was slightly distinct when using another amyloid antibody (Fig. S3a,b). Intra-neuronal accumulation of amyloid and reduced synaptophysin immuno-reactivity was also observed in the hippocampus of another *Bmi1*$^{+/-}$ mouse (Fig. S3c,d). This revealed that old *Bmi1*$^{+/-}$ mice could spontaneously develop a disorder sharing some similarities with AD.

Bmi1 hemi-deficiency cooperates with APP-mediated neurodegeneration. Genetic cooperation between two mutations can be indicative of protein-protein interactions and/or convergent molecular pathways. To test if *Bmi1* hemi-deficiency could cooperate or not with FAD-associated mutations, we crossed *Bmi1*$^{+/-}$ mice with human *APP* transgenic mice. From this, we obtained 4 genotypes in F1: WT, *Bmi1*$^{+/-}$, *APP* and *APP/Bmi1*$^{+/-}$ mice. Kaplan Mayer analysis revealed that although both *Bmi1*$^{+/-}$ and *APP* mice had reduced maximal and median lifespan compared to WT littermates, the median lifespan of the *APP/Bmi1*$^{+/-}$ animals was further reduced by ~45% when compared to *APP* mice (Fig. 4a). The difference was significant despite the limitations of using a small data set. When analyzed by IHC at 6 months of age for p-Tau, we observed almost no signal in WT and *Bmi1*$^{+/-}$ mice, and weak immunoreactivity in *APP* mice. Notably, while neuronal accumulation of p-Tau was rare in *APP* mice and absent in *Bmi1*$^{+/-}$ mice, it was relatively robust and abundant in *APP/Bmi1*$^{+/-}$ mice (Fig. 4b,c). Similarly, although intra-neuronal amyloid immunoreactivity was weak to absent in *Bmi1*$^{+/-}$ mice at this young age, it was easily detectable in *APP* mice (Fig. 4d). Notably, the number of immunoreactive neurons and the signal strength for amyloid were greatly enhanced in *APP/Bmi1*$^{+/-}$ mice (Fig. 4d,e). Other parameters of disease severity were also enhanced in *APP/Bmi1*$^{+/-}$ mice, including reactive gliosis, microglia activation, synaptic atrophy and apoptosis (Fig. S4). Taken together, these results suggested genetic cooperation between *Bmi1* hemi-deficiency and mutant *APP* in neurodegeneration.

Loss of heterochromatin and DDR in cortical neurons of *Bmi1*$^{+/-}$ mice. BMI1 is required for heterochromatin maintenance in somatic cells[19]. We thus investigated if the AD-like phenotype of *Bmi1*$^{+/-}$ mice was also associated with heterochromatin alterations. By Western blot analysis on whole cortices, we observed reduced levels for histones H2A^{ub} and H3K9^{me3} in 15 month-old *Bmi1*$^{+/-}$ mice (Fig. 5a). To test if this correlated with an abnormal chromatin structure, we performed IHC. We observed that in contrast to WT mice, where 2-3 large chromocenters/neuron were present, chromocenters of *Bmi1*$^{+/-}$ neurons were smaller and more numerous, suggesting heterochromatin de-nucleation (Fig. 5b). Consistently, this was associated with transcriptional de-repression of satellite repeats, which are normally silenced by heterochromatin (Fig. 5c)[38–40]. One prediction of heterochromatin de-nucleation is activation of a DDR. The *Ataxia Telengectasia Mutated* (ATM) and *Ataxia Telengectasia and RAD-3 Related* (ATR) kinases work at the apex of the DDR[41–43]. Using immunoblot, we observed accumulation of phospho-ATM at Serine 1981 (p-ATM) and phospho-ATR at Serine 428 (p-ATR) in the cortices of *Bmi1*$^{+/-}$ mice (Fig. 5d). The histone mark γH2AX is present at sites of DNA damage[44]. We observed nuclear accumulation of γH2AX in cortical neurons only in *Bmi1*$^{+/-}$ mice, suggesting that the DDR occurs primarily in neurons (Fig. 5e).

To localize DNA damage on the genome, we performed chromatin immuno-precipitation (ChIP) experiments on whole cortices. As expected, Bmi1, H3K9^{me3} and H2A^{ub} were enriched at genomic repeats (Minor and Major satellite repeats and LINE elements) and at the *Hoxa7* locus in WT mice and were depleted in *Bmi1*$^{+/-}$ mice (Fig. 5f)[19]. In contrast, while p-ATR and γH2AX were nearly absent in WT mice, they were preferentially and highly enriched at genomic repeats in *Bmi1*$^{+/-}$ mice (Fig. 5f).

To test if heterochromatin anomalies were present before disease onset, we performed immuno-blot, IHC and confocal microscopy on cortices from 2-3 month-old WT and *Bmi1*$^{+/-}$ mice. This revealed reduced H3K9^{me3} levels and fragmented H3K9^{me3} immunostaining in *Bmi1*$^{+/-}$ pyramidal neurons when compared to WT (Figs 5g and S5a,b). Reduced H3K9^{me3} levels were also present in cultured embryonic day 18.5 neurons, as measured using ChIP analyses (Fig. S5c). Despite the presence of chromatin anomalies at the earliest step of neurogenesis, we failed to detect a DDR or the expression of AD-related markers in 3 month-old *Bmi1*$^{+/-}$ mice using immuno-blot (Fig. S5d). Mild accumulation of p53 was however noticed in *Bmi1*$^{+/-}$ mice (Fig. S5d). This suggested that in *Bmi1*$^{+/-}$ mice, loss of heterochromatin is a very early event in the disease process.

Amyloid and Tau pathologies can be mitigate by inhibition of the DDR. *Bmi1*-null (*Bmi1*$^{-/-}$) mice are smaller than normal and die around post-natal day 30[12]. Yet, we could find cortical neurons immuno-reactive for p-Tau and amyloid in 25 day-old *Bmi1*$^{-/-}$ mice (Fig. 6a,b). To gain insight into the disease mechanism, we generated cohorts of 25 day-old *Bmi1*$^{-/-}$ mice carrying null mutations in *p53* (*Bmi1*$^{-/-}$/*p53*$^{-/-}$) or *Chk2* (*Bmi1*$^{-/-}$/*Chk2*$^{-/-}$). *Chk2* encodes a kinase activated by ATM, and both ATM and Chk2 can stimulate p53 activity[45,46]. Notably, we found that p-Tau and amyloid immuno-reactivity were both mitigated in *Bmi1*$^{-/-}$/*p53*$^{-/-}$and in *Bmi1*$^{-/-}$/*Chk2*$^{-/-}$ mice (Fig. 6a,b).

To understand the basis of p53 activation, we cultured WT, *Bmi1*$^{-/-}$ and *Bmi1*$^{+/-}$ embryonic day 18.5 cortical neurons to test if free radical-mediated DNA damage was responsible for p53 accumulation[18,47,48]. Using N-Acetyl Cysteine (NAC), a free radical scavenger, we found that reducing free radicals could slightly improve p53 levels in

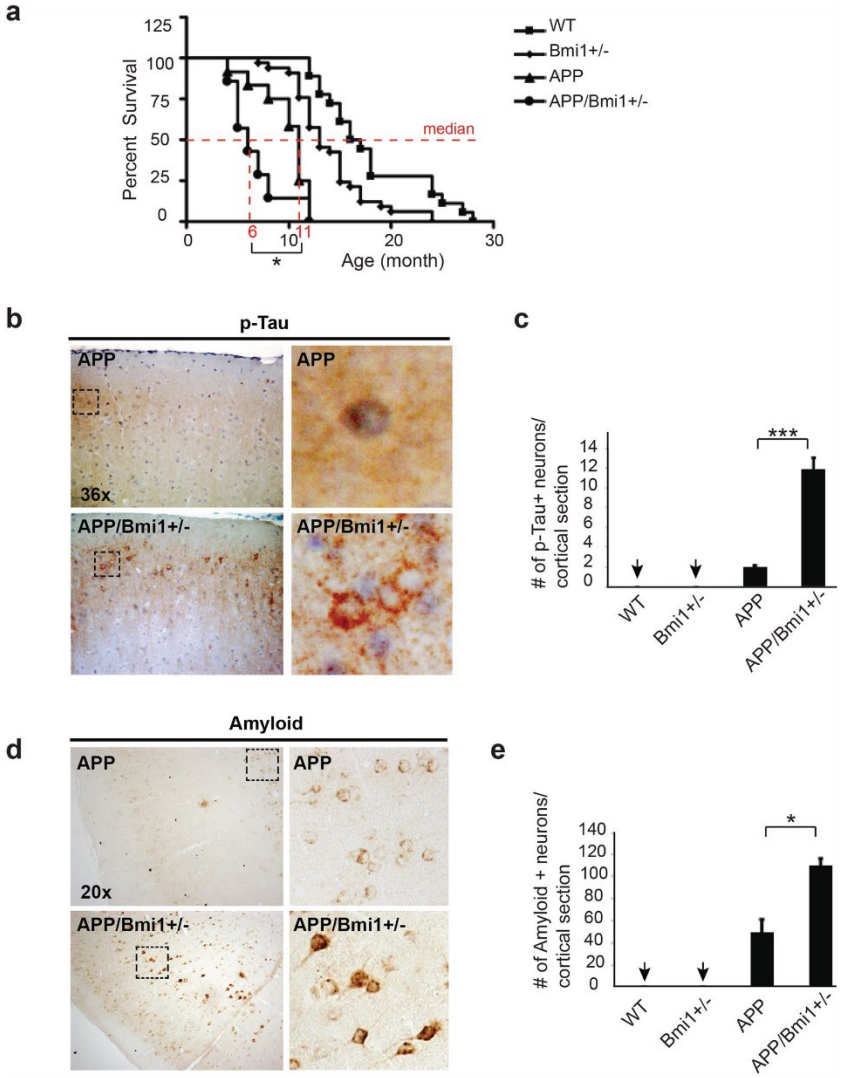


Figure 4. *APP/Bmi1*$^{+/-}$ mice show genetic cooperation in disease onset and severity. (**a**) Kaplan-Meyer representation of the survival curves of WT (n = 11), *Bmi1*$^{+/-}$ (n = 13), *APP* (n = 6) and *APP/Bmi1*$^{+/-}$ (n = 6) mice. (**b**) IHC staining for p-Tau (PHF1) on cortical sections of 6 month-old WT (n = 3), *Bmi1*$^{+/-}$ (n = 3), *APP* (n = 3),and *APP/Bmi1*$^{+/-}$ (n = 3) mice. Scale bar: 20μm. Scale bar in the inset: 8μm. (**c**) Quantification of p-Tau-positive neurons shown in (**b**). (**d**) IHC for amyloid (DE2B4) on cortical sections of 6-month old WT (n = 3), *Bmi1*$^{+/-}$ (n = 3), *APP* (n = 3), and *APP/Bmi1*$^{+/-}$ (n = 3) mice. Scale bar: 20μm. (**e**) Quantification of amyloid-positive cells shown in (**d**). Values are mean ± SEM. *$P < 0.05$; (**) < 0.01; (***) < 0.001; Student's t-test.

Bmi1$^{-/-}$ and *Bmi1*$^{+/-}$ neurons (Fig. 6c). The p19^{Arf} protein can stabilize p53, and p19^{Arf} is up regulated in *Bmi1*$^{-/-}$ neurons[18,49]. However, co-deletion of *Bmi1* and *Ink4a* (encoding for p16^{Ink4a} and p19^{Arf}) could not prevent p53 accumulation (Fig. 6d). In response to DNA damage, the ATM and ATR kinases can target p53 for phosphorylation[45,50,51]. We found that inhibition of ATM alone could not prevent p53 accumulation, but that inhibition of ATM and ATR largely mitigated p53 accumulation in *Bmi1*$^{-/-}$ neurons (Fig. 6e). Residual p53 levels were however observed. We also observed the presence of p-ATM and phospho-Chk1 (a downstream target of ATR) in *Bmi1*$^{-/-}$ neurons and found that ATM/ATRi was sufficient to prevent their accumulation (Fig. 6f). Remarkably,

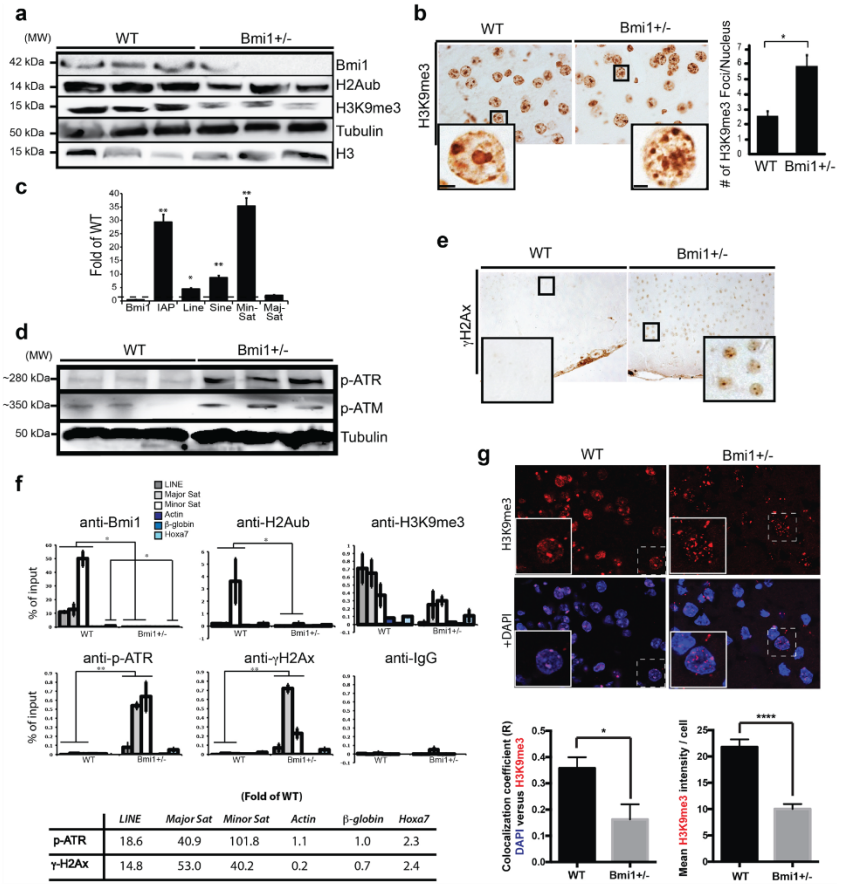


Figure 5. Loss of heterochromatin and DDR in cortical neurons of *Bmi1*$^{+/-}$ mice. (**a**) Western blot analysis of cortical extracts from old WT (n = 6) and *Bmi1*$^{+/-}$ (n = 6) mice. (**b**) IHC on cortical sections from old WT (n = 3) and *Bmi1*$^{+/-}$ (n = 3) mice with quantification of the number of H3K9^{me3} foci per nucleus (6 sections/sample were counted). Scale bar in the inset: 3µm. (**c**) Real-time RT-PCR analysis for the expression of satellite repeats in cortices from WT (n = 3) and *Bmi1*$^{+/-}$ (n = 3) mice. (**d**) Western blot analysis of cortical extracts from old WT (n = 3) and *Bmi1*$^{+/-}$ (n = 3) mice. (**e**) IHC for γH2AX on cortical sections of old WT and *Bmi1*$^{+/-}$ mice. (**f**) ChIP analyses were performed on frontal cortex homogenates from 15 month-old WT (n = 3) and *Bmi1*$^{+/-}$ (n = 3) mice. Quantitative PCR was performed in triplicate for each DNA sequence. All data are presented as fold of input. Bmi1 and H2A^{ub} reduction ($p < 0.05$), and p-ATR and γH2AX accumulation ($p < 0.01$) in *Bmi1*$^{+/-}$ mice are significant (Two-way ANOVA analysis). (bottom panel) Fold differences for γH2AX and p-ATR accumulation in *Bmi1*$^{+/-}$ mice were measured relative to WT levels for each DNA sequences. (**g**) Cortical slices from 45-day old mice. Areas showed are located in layers 2-3 of the frontal cortex. In WT samples, large pyramidal neurons were strongly labeled by the H3K9^{me3} antibody and H3K9^{me3} co-localized with DAPI. In *Bmi1*$^{+/-}$ samples, large pyramidal neurons were weakly labeled by the H3K9^{me3} antibody and the H3K9me3 signal was fragmented. Bottom panels: Quantification of the images showed in (**f**). Values are mean ± SEM. *$P < 0.05$; **$P < 0.01$; ****$P < 0.0001$; Student's t-test.

p-Tau accumulation in *Bmi1*$^{-/-}$ neurons could be also prevented after exposure to ATM/ATRi (Fig. 6f). These results suggested a theoretical model where AD-related neuronal pathologies observed in *Bmi1*-deficient mice are triggered in part by a ATM/ATR-driven DDR working upstream of p53, but where Bmi1/Ring1 ubiquitin ligase activity and Bmi1-mediated ROS repression are also likely required to prevent p53 accumulation (Fig. 6g)[52].

Loss of heterochromatin and DDR in AD brains. We next investigated the possible relevance of our findings in the context of AD. We first confirmed the robust expression of BMI1 in cortical neurons from aged controls and BMI1 depletion in neurons from AD patients (Fig. 7a). To test for possible heterochromatin

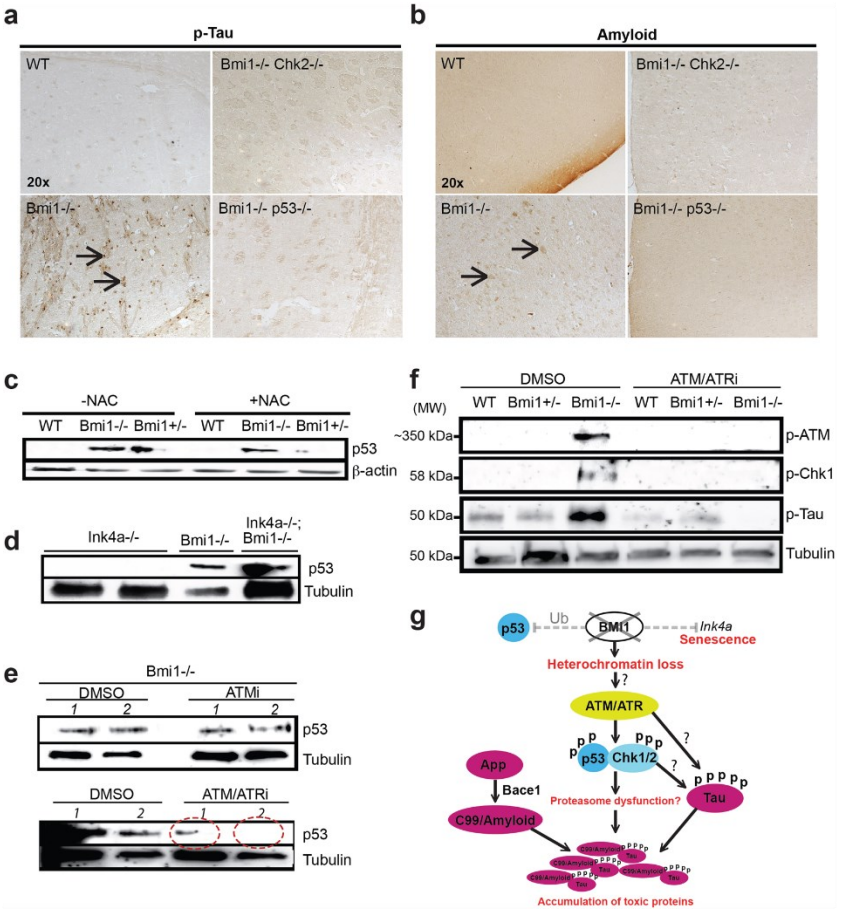
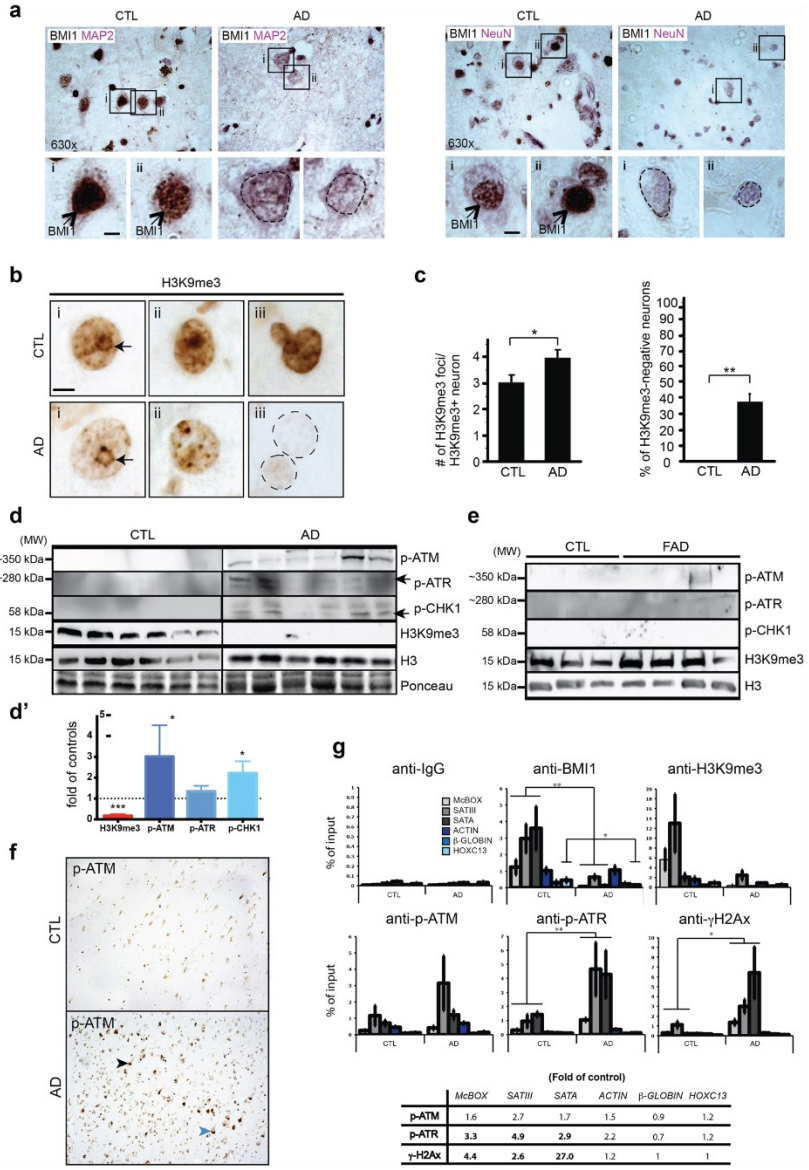


Figure 6. Blocking the DDR mitigate the amyloid and Tau phenotype in *Bmi1*-null neurons. (**a**,**b**) IHC performed on cortical sections from 25 day-old mice. (**a**) Arrows indicate p-Tau-positive neurons in *Bmi1*$^{-/-}$ mice. (**b**) Arrows indicate amyloid-positive neurons in *Bmi1*$^{-/-}$ mice. These anomalies were largely prevented in the double-mutant mice. (**c**–**f**) Immunoblot were performed on embryonic day 18.5 mouse cortical neurons cultured for 7 days *in vitro*. (**c**–**d**) Addition of NAC or co-deletion of *Ink4a* could not prevent p53 accumulation in *Bmi1*-deficient neurons. Note the presence of p53 also in *Bmi1*$^{+/-}$ neurons. NAC was added daily to the cultures for a period of 7 days. (**e**,**f**) DMSO, ATMi or ATM/ATRi were added to the cultures 16 hours prior to protein extraction. (**e**) ATM/ATRi largely prevented p53 accumulation in *Bmi1*$^{-/-}$ neurons. Numbers *1* and *2* in (**e**) indicate two independent biological replicates. Note residual p53 expression in ATM/ATRi-treated neurons (circles with dash lines). (**f**) Addition of ATM/ATRi could prevent p-ATM, p-Chk1 and p-Tau accumulation in *Bmi1*$^{-/-}$ neurons. (**g**) Hypothetical model to explain the impact of *Bmi1* inactivation in mouse neurons. Following loss of *Bmi1*, up-regulation of the *Ink4a* locus induces neuronal senescence. On the other hand, genomic instability at the heterochromatin induces a chronic DDR. The p53 protein is then stabilized by ATR/ATM-mediated kinase activity. The accumulation of p53 may be also exacerbated by the loss of Bmi1/Ring1a-mediated ubiquitin ligase and antioxidant activities. The activated ATM/Chk2, ATR/Chk1 and p53 pathways then promote Tau phosphorylation and accumulation of the C99 fragment. Elevated Bace1 levels also contribute to over-production of the C99 fragment. Accumulation of these toxic proteins contributes to neurodegeneration.

anomalies, we performed IHC on frontal cortex sections[53]. In control samples, all neurons were labeled with H3K9^{me3} and presented ~3 heterochromatic foci (or chromocenters)/neuron (Fig. 7b,c). In AD samples, ~38% of all pyramidal neurons showed highly reduced or absent H3K9^{me3} nuclear staining, and the neuron's chromocenters were smaller and more numerous (Fig. 7b,c). To test the possibility of genomic instability, we analyzed



	McBOX	SATIII	SATA	ACTIN	β-GLOBIN	HOXC13
p-ATM	1.6	2.7	1.7	1.5	0.9	1.2
p-ATR	**3.3**	**4.9**	**2.9**	2.2	0.7	1.2
γ-H2Ax	**4.4**	**2.6**	**27.0**	1.2	1	1

Figure 7. Loss of heterochromatin and DDR in AD brains. (**a**) IHC staining for BMI1 (brown staining-black arrows) and MAP2 or NeuN (violet staining) on paraffin sections from the frontal cortex of age-match control (n = 2) and AD (n = 2) patients. Weak nuclear staining for BMI1 (dash lines) was observed in AD neurons. Scale bar: 3μm. (**b**) IHC staining for $H3K9^{me3}$ on frontal cortex sections of control (n = 5) and AD (n = 5) brains: (i) reduced peri-nucleolar heterochromatin in AD (arrows), (ii) chromocenters de-condensation in AD, and (iii) loss of $H3K9^{me3}$ immunoreactivity in AD neurons (dash lines). Scale bar: 5μm. (**c**) Quantification of data in (b) showing the number of $H3K9^{me3}$ foci/$H3K9^{me3}$-positive neuron. $H3K9^{me3}$-negative neurons were excluded from the analysis. Quantification of the number of $H3K9^{me3}$-negative neurons, which were not found in aged controls. 6 full-fields at 630 magnification/sample were counted. Values are mean ± SEM. $*P < 0.05$; (**) < 0.01; Student's t-test. (**d**) Immunoblot using extracts from the frontal cortex of age-match control and AD patients. (**d'**) Quantification of the results showed in (**d**). (**e**) Immunoblot using extracts from the hippocampus of age-match control and FAD patients. (**f**) IHC staining for p-ATM on paraffin sections from

the frontal cortex of age-match control (n = 2) and AD (n = 2) patients. p-ATM is predominant in AD neurons and accumulates in both cytosolic (blue arrowhead) and nuclear (black arrowhead) compartments. (**g**) ChIP experiments were performed on frontal cortex extracts of age-match control (n = 6) and AD brains (n = 6). All data are represented as *fold of input*. (A) Note BMI1 and H3K9^{me3} enrichment at *McBOX*, *SATIII* and *SATA* in control samples. BMI1 was also found at *HOXC13*. BMI1 and H3K9^{me3} were depleted at all loci in AD samples (bottom). Fold differences between control and AD samples for p-ATM, p-ATR and γH2AX accumulation. Values are mean ± SEM. *$P < 0.05$; **$P < 0.01$; ***$P < 0.001$; Two way-ANOVA test was performed for multiple gene analysis and Student's t-test for single gene analysis.

the expression of DDR proteins. By Western blot analysis, we observed accumulation of p-ATM, p-ATR and p-CHK1 in AD brains, but not in age-match controls (Fig. 7d,d'). The presence of a DDR in AD brains also correlated with reduced H3K9^{me3} levels (Fig. 7d). In contrast, accumulation of DDR proteins was not observed in FAD brains, with the exception of one sample presenting p-ATM accumulation (Fig. 7e). Likewise, H3K9^{me3} levels were unaffected in FAD brains (Fig. 7e). Using IHC, we also confirmed that DDR in AD brains occurred primarily in neurons (Fig. 7f).

To establish where the DDR was located on the genome, we performed ChIP experiments using frontal cortex samples. When compared to controls, we found significant enrichment for p-ATR and γH2AX at genomic repeats (*SATIII*, *SATA* and *McBOX*) in AD brains (Fig. 7g). In contrast, there was no enrichment at *HOXC13*, β-*GLOBIN* or *ACTIN* (Fig. 7g). Although there was a general trend for p-ATM enrichment at genomic repeats in AD samples, it was not significant. BMI1 and H3K9^{me3} were also enriched at *HOXC13* and at genomic repeats in control brains and were depleted in AD brains (Fig. 7g). Taken together, these findings revealed remarkable similarities between the neuronal genomic instability phenotype of *Bmi1*$^{+/-}$ mice and AD patients.

Discussion

We have found that *Bmi1*$^{+/-}$ mice develop with age a neurological and progeroid syndrome characterized by alopecia, weight loss, abnormal paw clasping reflex and neuronal senescence. Memory deficit as measured by their performance in the Morris water maze and Barnes test correlated with impairment in LTP formation as well as with synaptic atrophy and neuronal loss. Pathological brain anomalies such as p-Tau and amyloid accumulation were also found in old *Bmi1*$^{+/-}$ mice. Remarkably, when compared to *Bmi1*$^{+/-}$ or *APP* mice alone, *Bmi1*$^{+/-}$/*APP* mice showed significantly shorter lifespan and increased disease severity. Genetic and cell culture studies further revealed that inhibition of the DDR and/or p53 could mitigate the accumulation of p-Tau and amyloid in neurons from *Bmi1*-null mice, suggesting that the neuronal response to genomic instability plays a role in the disease process. At last, loss of heterochromatin in neurons and accumulation of DNA damage at genomic repeats were found in *Bmi1*$^{+/-}$ mice and AD brains.

One possible interpretation of our data is that *Bmi1*$^{+/-}$ mice age faster and thus present age-related neurological anomalies earlier than WT mice. This possibility is supported by our observation that some pathological marks, including p-Tau (using PHF1, which recognizes an epitope around Ser396 and Ser404) and SA β-galactosidase activity, although much less abundant, were also detected in the brains of old WT mice. However, other pathological marks, such as p-Tau (using AT-8, which recognizes Ser202), p-JNK and amyloid (C99) accumulation were not detected even in 24 month-old WT mice, suggesting that *Bmi1* hemi-deficiency represents a pathological brain aging process that is distinct from end-stage physiological brain aging.

Using the *Bmi1*$^{+/-}$ and *Bmi1*$^{-/-}$ mouse models, we observed activation of DDR proteins and of p53 in neurons. This is consistent with previous work showing that *Bmi*1 is important to maintain genomic stability[15–18,54]. Interestingly, over-expression of *Chk1* or *Chk*2 was shown to exacerbate Tau toxicity in a *Drosophila* model of neurodegeneration by specific phosphorylation of human Tau at Ser262[55]. Likewise, JNK activation may promote Tau phosphorylation[32]. Thus, several kinases potentially targeting Tau are activated in *Bmi1*$^{+/-}$ mice. It however remains to be investigated by which mechanisms Bace1 level is increased and why the amyloid metabolism is affected in *Bmi1*$^{+/-}$ mice. Our results showed that amyloid accumulation was mitigated after *p53* deletion and we recently reported that p53 activation in human neurons might perturb the clearance of misfolded proteins by interfering with proteasome function[20]. Thus, a plausible mechanism to explain our results is that activated p53 also impairs proteasome function in mouse neurons. Further work is however needed to validate this.

The heterochromatin island hypothesis of aging stipulates that epigenome instability at the constitutive heterochromatin is a driving force of cellular aging[56,57]. This hypothesis may be particularly relevant in post-mitotic neurons, where cell division associated with telomere attrition is irrelevant to the cellular aging process. Notably, inactivation of chromatin remodeling proteins, such as those of the NURD complex, results in DDR and in cellular senescence that are preceded by heterochromatin alterations[58]. Furthermore, modifications of constitutive heterochromatin in response to an acute stress can result in epigenome instability and perturbation of gene expression program, which are hallmarks of cellular aging[59]. In mice, deletion of the SUV39h1/h2 methyltransferases leads to reduced viability, loss of H3K9^{me3} and genomic instability[60]. In *Drosophila*, perturbation of H3K9^{me2} levels through inactivation of Su(var)3-9 results in genomic instability and constitutive DDR at the heterochromatin, in both somatic and germ-line cells[61]. We showed here that heterochromatin alterations in *Bmi1*$^{+/-}$ mice are present at the time of neurogenesis and thus before activation of DDR proteins and neurodegeneration. Notably, this correlated with preferential accumulation of DNA damage (as marked by γH2Ax) at repetitive DNA sequences. Similarly, we found that DNA damage accumulation was predominant at genomic repeats in AD brains. Exactly why loss of heterochromatin compaction strongly correlates with DNA damage accumulation at genomic repeats is unknown. It is possible that repetitive DNA sequences are intrinsically unstable upon loss of heterochromatin compaction and thus prone to *de novo* DNA damage formation. It is also possible that the

efficiency of DNA repair is reduced upon BMI1 deficiency or that recognition of the H3K9^{me3} mark is important to activate the process of DNA repair at these specific regions[15,16,62]. It remains however to be tested whether heterochromatin anomalies observed in *Bmi1*$^{+/-}$ mice are sufficient to induce the DDR and neurodegeneration.

Interstitial amyloid plaques and Tau tangles represent the classical pathological hallmarks of AD. While synaptic atrophy, neuronal loss and behavioral anomalies are present in 15-month old *Bmi1*$^{+/-}$ mice, the animals do not develop Tau tangles or amyloid plaques (with the exception of one animal). Rather, old *Bmi1*$^{+/-}$ mice present intra-neuronal accumulation of amyloid and show moderate levels of p-Tau deposits in the neuronal cell body and the axon. The absence of amyloid plaques formation in *Bmi1*$^{+/-}$ mice is not surprising since the mouse amyloid peptide is known to be less prone to aggregation than the human one[63]. Also, transgenic mice over-expressing wild-type human *APP* rarely develop amyloid plaques, in contrast with those over-expressing *APP* with FAD-associated mutations[64]. Notably, massive neurodegeneration was reported in *APP/PSEN1* transgenic mice (*APP(SL)PS1KI* mice), which only develop intra-neuronal accumulation of amyloid[65]. Thus, intra-neuronal accumulation of oligomeric amyloid species, including the C99 fragment, may be sufficient to induce neurodegeneration[66,67]. Nevertheless, the absence of amyloid plaques and Tau tangles as well as the late onset neuronal disease should be taken into account before using *Bmi1*$^{+/-}$ mice as a model of AD.

In conclusion, we demonstrated that loss of one *Bmi1* allele in mice results in age-related neurodegeneration sharing some similarities with AD. *Bmi1*$^{+/-}$ mice may thus represent an interesting animal model to identify new pathogenic mechanisms related to AD.

Data Availability Statement

The authors declare that they will make available all data presented in this manuscript.

References

1. Savva, G. M. *et al.* Age, neuropathology, and dementia. *N. Engl. J. Med.* **360**, 2302–2309 (2009).
2. Kanekiyo, T. *et al.* ApoE and Abeta in Alzheimer's disease: accidental encounters or partners? *Neuron* **81**, 740–754 (2014).
3. Blennow, K. *et al.* Alzheimer's disease. *Lancet* **368**, 387–403 (2006).
4. Chapman, P. F. *et al.* Impaired synaptic plasticity and learning in aged amyloid precursor protein transgenic mice. *Nat. Neurosci.* **2**, 271–276 (1999).
5. Galvan, V. *et al.* Reversal of Alzheimer's-like pathology and behavior in human APP transgenic mice by mutation of Asp664. *Proc. Natl. Acad. Sci. USA* **103**, 7130–7135 (2006).
6. Tomiyama, T. *et al.* A mouse model of amyloid beta oligomers: their contribution to synaptic alteration, abnormal tau phosphorylation, glial activation, and neuronal loss *in vivo*. *J. Neurosci.* **30**, 4845–4856 (2010).
7. Fodor, B. D. *et al.* Mammalian Su(var) genes in chromatin control. *Annu. Rev. Cell Dev. Biol.* **26**, 471–501 (2010).
8. Buchwald, G. *et al.* Structure and E3-ligase activity of the Ring-Ring complex of polycomb proteins Bmi1 and Ring1b. *EMBO J.* **25**, 2465–2474 (2006).
9. Li, Z. *et al.* Structure of a Bmi-1-Ring1B polycomb group ubiquitin ligase complex. *J. Biol. Chem.* **281**, 20643–20649 (2006).
10. Satijn, D. P. *et al.* RING1 is associated with the polycomb group protein complex and acts as a transcriptional repressor. *Mol. Cell. Biol.* **17**, 4105–4113 (1997).
11. Wang, H. *et al.* Role of histone H2A ubiquitination in Polycomb silencing. *Nature* **431**, 873–878 (2004).
12. Jacobs, J. J. *et al.* The oncogene and Polycomb-group gene bmi-1 regulates cell proliferation and senescence through the ink4a locus. *Nature* **397**, 164–168 (1999).
13. Molofsky, A. V. *et al.* Bmi-1 dependence distinguishes neural stem cell self-renewal from progenitor proliferation. *Nature* **425**, 962–967 (2003).
14. van der Lugt, N. M. *et al.* Posterior transformation, neurological abnormalities, and severe hematopoietic defects in mice with a targeted deletion of the bmi-1 proto-oncogene. *Genes Dev.* **8**, 757–769 (1994).
15. Facchino, S. *et al.* BMI1 confers radioresistance to normal and cancerous neural stem cells through recruitment of the DNA damage response machinery. *J. Neurosci.* **30**, 10096–10111 (2010).
16. Ismail, I. H. *et al.* BMI1-mediated histone ubiquitylation promotes DNA double-strand break repair. *J. Cell Biol.* **191**, 45–60 (2010).
17. Liu, J. *et al.* Bmi1 regulates mitochondrial function and the DNA damage response pathway. *Nature* **459**, 387–392 (2009).
18. Chatoo, W. *et al.* The polycomb group gene Bmi1 regulates antioxidant defenses in neurons by repressing p53 pro-oxidant activity. *J. Neurosci.* **29**, 529–542 (2009).
19. Abdouh, M. *et al.* The Polycomb Repressive Complex 1 Protein BMI1 is Required for Constitutive Heterochromatin Formation and Silencing in Mammalian Somatic Cells. *J. Biol. Chem.* (2015).
20. Flamier, A. *et al.* Modeling late-onset sporadic Alzheimer's disease through BMI1 deficiency. *Cell Reports* **23**, 2653–2666 (2018).
21. Bornemann, K. D. & Staufenbiel, M. Transgenic mouse models of Alzheimer's disease. *Ann. N. Y. Acad. Sci.* **908**, 260–266 (2000).
22. Donehower, L. A. *et al.* Mice deficient for p53 are developmentally normal but susceptible to spontaneous tumours. *Nature* **356**, 215–221 (1992).
23. Barabino, A. *et al.* Loss of Bmi1 causes anomalies in retinal development and degeneration of cone photoreceptors. *Development* **143**, 1571–1584 (2016).
24. Kawarabayashi, T. *et al.* Age-dependent changes in brain, CSF, and plasma amyloid (beta) protein in the Tg2576 transgenic mouse model of Alzheimer's disease. *J. Neurosci.* **21**, 372–381 (2001).
25. Dimri, G. P. *et al.* A biomarker that identifies senescent human cells in culture and in aging skin *in vivo*. *Proc. Natl. Acad. Sci. USA* **92**, 9363–9367 (1995).
26. Baker, D. J. *et al.* Early aging-associated phenotypes in Bub3/Rae1 haploinsufficient mice. *J. Cell Biol.* **172**, 529–540 (2006).
27. Kondratov, R. V. *et al.* Early aging and age-related pathologies in mice deficient in BMAL1, the core componentof the circadian clock. *Genes Dev.* **20**, 1868–1873 (2006).
28. Ruzankina, Y. *et al.* Deletion of the developmentally essential gene ATR in adult mice leads to age-related phenotypes and stem cell loss. *Cell Stem Cell* **1**, 113–126 (2007).
29. Trifunovic, A. *et al.* Premature ageing in mice expressing defective mitochondrial DNA polymerase. *Nature* **429**, 417–423 (2004).
30. Tyner, S. D. *et al.* p53 mutant mice that display early ageing-associated phenotypes. *Nature* **415**, 45–53 (2002).
31. Lalonde, R. & Strazielle, C. Motor performance and regional brain metabolism of spontaneous murine mutations with cerebellar atrophy. *Behav. Brain Res.* **125**, 103–108 (2001).
32. Wetzel, M. K. *et al.* p73 regulates neurodegeneration and phospho-tau accumulation during aging and Alzheimer's disease. *Neuron* **59**, 708–721 (2008).
33. Krishnamurthy, J. *et al.* Ink4a/Arf expression is a biomarker of aging. *J. Clin. Invest.* **114**, 1299–1307 (2004).
34. Simon, P. *et al.* Thigmotaxis as an index of anxiety in mice. *Influence of dopaminergic transmissions. Behav. Brain Res.* **61**, 59–64 (1994).

35. Bliss, T. V. & Collingridge, G. L. A synaptic model of memory: long-term potentiation in the hippocampus. *Nature* **361**, 31–39 (1993).
36. Sclip, A. *et al*. c-Jun N-terminal kinase regulates soluble Abeta oligomers and cognitive impairment in AD mouse model. *J. Biol. Chem.* **286**, 43871–43880 (2011).
37. Yang, J. *et al*. Pathological Axonal Death through a MAPK Cascade that Triggers a Local Energy Deficit. *Cell* **160**, 161–176 (2015).
38. Kim, M. *et al*. Dnmt1 deficiency leads to enhanced microsatellite instability in mouse embryonic stem cells. *Nucleic Acids Res* **32**, 5742–5749 (2004).
39. Lehnertz, B. *et al*. Suv39h-mediated histone H3 lysine 9 methylation directs DNA methylation to major satellite repeats at pericentric heterochromatin. *Curr. Biol.* **13**, 1192–1200 (2003).
40. Zhu, Q. *et al*. BRCA1 tumour suppression occurs via heterochromatin-mediated silencing. *Nature* **477**, 179–184 (2011).
41. Cimprich, K. A. & Cortez, D. ATR: an essential regulator of genome integrity. *Nat Rev Mol Cell Biol* **9**, 616–627 (2008).
42. Lavin, M. F. ATM and the Mre11 complex combine to recognize and signal DNA double-strand breaks. *Oncogene* **26**, 7749–7758 (2007).
43. Smith, J. *et al*. The ATM-Chk2 and ATR-Chk1 pathways in DNA damage signaling and cancer. *Adv. Cancer Res.* **108**, 73–112 (2010).
44. Rogakou, E. P. *et al*. DNA double-stranded breaks induce histone H2AX phosphorylation on serine 139. *J. Biol. Chem.* **273**, 5858–5868 (1998).
45. Keramaris, E. *et al*. Ataxia telangiectasia-mutated protein can regulate p53 and neuronal death independent of Chk2 in response to DNA damage. *J. Biol. Chem.* **278**, 37782–37789 (2003).
46. Chaturvedi, P. *et al*. Mammalian Chk2 is a downstream effector of the ATM-dependent DNA damage checkpoint pathway. *Oncogene* **18**, 4047–4054 (1999).
47. Brady, C. A. *et al*. Distinct p53 transcriptional programs dictate acute DNA-damage responses and tumor suppression. *Cell* **145**, 571–583 (2011).
48. Zhang, X. P. *et al*. Two-phase dynamics of p53 in the DNA damage response. *Proc. Natl. Acad. Sci. USA* **108**, 8990–8995 (2011).
49. Mellert, H. *et al*. The ARF/oncogene pathway activates p53 acetylation within the DNA binding domain. *Cell Cycle* **6**, 1304–1306 (2007).
50. Lakin, N. D. *et al*. The ataxia-telangiectasia related protein ATR mediates DNA-dependent phosphorylation of p53. *Oncogene* **18**, 3989–3995 (1999).
51. Tibbetts, R. S. *et al*. A role for ATR in the DNA damage-induced phosphorylation of p53. *Genes Dev.* **13**, 152–157 (1999).
52. Calao, M. *et al*. Direct effects of Bmi1 on p53 protein stability inactivates oncoprotein stress responses in embryonal cancer precursor cells at tumor initiation. *Oncogene* **32**, 3616–3626 (2013).
53. Frost, B. *et al*. Tau promotes neurodegeneration through global chromatin relaxation. *Nat. Neurosci.* **17**, 357–366 (2014).
54. Chagraoui, J. *et al*. An anticlastogenic function for the Polycomb Group gene Bmi1. *Proc. Natl. Acad. Sci. USA* **108**, 5284–5289 (2011).
55. Iijima-Ando, K. *et al*. A DNA damage-activated checkpoint kinase phosphorylates tau and enhances tau-induced neurodegeneration. *Hum. Mol. Genet.* **19**, 1930–1938 (2010).
56. Imai, S. & Kitano, H. Heterochromatin islands and their dynamic reorganization: a hypothesis for three distinctive features of cellular aging. *Exp. Gerontol.* **33**, 555–570 (1998).
57. Villeponteau, B. The heterochromatin loss model of aging. *Exp. Gerontol.* **32**, 383–394 (1997).
58. Pegoraro, G. *et al*. Ageing-related chromatin defects through loss of the NURD complex. *Nat Cell Biol* **11**, 1261–1267 (2009).
59. Oberdoerffer, P. *et al*. SIRT1 redistribution on chromatin promotes genomic stability but alters gene expression during aging. *Cell* **135**, 907–918 (2008).
60. Peters, A. H. *et al*. Loss of the Suv39h histone methyltransferases impairs mammalian heterochromatin and genome stability. *Cell* **107**, 323–337 (2001).
61. Peng, J. C. & Karpen, G. H. Heterochromatic genome stability requires regulators of histone H3 K9 methylation. *PLoS Genet* **5**, e1000435 (2009).
62. Sun, Y. *et al*. Histone H3 methylation links DNA damage detection to activation of the tumour suppressor Tip60. *Nat Cell Biol* **11**, 1376–1382 (2009).
63. De Strooper, B. *et al*. Production of intracellular amyloid-containing fragments in hippocampal neurons expressing human amyloid precursor protein and protection against amyloidogenesis by subtle amino acid substitutions in the rodent sequence. *EMBO J.* **14**, 4932–4938 (1995).
64. Sasaguri, H. *et al*. APP mouse models for Alzheimer's disease preclinical studies. *EMBO J.* **36**, 2473–2487 (2017).
65. Casas, C. *et al*. Massive CA1/2 neuronal loss with intraneuronal and N-terminal truncated Abeta42 accumulation in a novel Alzheimer transgenic model. *Am. J. Pathol.* **165**, 1289–1300 (2004).
66. LaFerla, F. M. *et al*. Intracellular amyloid-beta in Alzheimer's disease. *Nat Rev Neurosci* **8**, 499–509 (2007).
67. Lauritzen, I. *et al*. The beta-secretase-derived C-terminal fragment of betaAPP, C99, but not Abeta, is a key contributor to early intraneuronal lesions in triple-transgenic mouse hippocampus. *J. Neurosci.* **32**, 16243–16255a (2012).

Acknowledgements

We are grateful to D. Cécyre and J. Prud'homme for the human brain samples, J. Rochford for help with the behavioral analysis, and to Drs F. Rodier, R. Kothary, G. Ferbeyre and D. Picketts for critical reading of the manuscript. This project was supported by grants from the Canadian Institutes of Health Research (374972) and Natural Science and Engineering Research Council of Canada (2017-05504). J.E.H. and R.H. held a fellowship from the University of Montreal Molecular Biology Program. G.B. was supported by a fellowship from the Fonds de Recherche en Santé du Québec.

Author Contributions

J.E.H., R.H., W.C., M.A., N.T., P.N. and Y.C.T. performed the experiments. G.B., J.E.H. and T.P.W. wrote the manuscript.

Additional Information

Supplementary information accompanies this paper at https://doi.org/10.1038/s41598-018-37444-3.

Competing Interests: G.B. is a co-founder of StemAxon™ and member of its scientific advisory board. The corporation was however not involved in this study.

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

8 Annex III: Loss of Bmi1 causes anomalies in retinal development and degeneration of cone photoreceptors

The second annex is an article, on which I have worked during my Ph.D., and was published on Development in 2016. In this study, we continued the work on BMI1 function as a neuroprotector but in the context of photoreceptors rather then in the context of cortical neurons. This work proved for the first time that photoreceptors lacking BMI1 undergo degeneration via RIP3-associated necroptosis.

© 2016. Published by The Company of Biologists Ltd | Development (2016) 143, 1571-1584 doi:10.1242/dev.125351

RESEARCH ARTICLE

Loss of Bmi1 causes anomalies in retinal development and degeneration of cone photoreceptors

Andrea Barabino[1,*], Vicky Plamondon[1,*], Mohamed Abdouh[1], Wassim Chatoo[1], Anthony Flamier[1], Roy Hanna[1], Shufeng Zhou[1], Noboru Motoyama[2], Marc Hébert[3], Joëlle Lavoie[3] and Gilbert Bernier[1,4,5,‡]

ABSTRACT

Retinal development occurs through the sequential but overlapping generation of six types of neuronal cells and one glial cell type. Of these, rod and cone photoreceptors represent the functional unit of light detection and phototransduction and are frequently affected in retinal degenerative diseases. During mouse development, the Polycomb group protein Bmi1 is expressed in immature retinal progenitors and differentiated retinal neurons, including cones. We show here that Bmi1 is required to prevent post natal degeneration of cone photoreceptors and bipolar neurons and that inactivation of Chk2 or p53 could improve but not overcome cone degeneration in *Bmi1*$^{-/-}$ mice. The retinal phenotype of *Bmi1*$^{-/-}$ mice was also characterized by loss of heterochromatin, activation of tandem repeats, oxidative stress and Rip3-associated necroptosis. In the human retina, BMI1 was preferentially expressed in cones at heterochromatic foci. BMI1 inactivation in human embryonic stem cells was compatible with retinal induction but impaired cone terminal differentiation. Despite this developmental arrest, BMI1-deficient cones recapitulated several anomalies observed in *Bmi1*$^{-/-}$ photoreceptors, such as loss of heterochromatin, activation of tandem repeats and induction of *p53*, revealing partly conserved biological functions between mouse and man.

KEY WORDS: Polycomb, Bmi1, Retina, Retinal development, Bipolar neuron, Cone, Photoreceptor, Degeneration, Mouse, Human, Embryonic stem cell

INTRODUCTION

The distinct competence of retinal progenitor cells (RPCs) to generate in a sequential order the diverse class of neurons and a single glial cell type during retinal development is thought to be modulated by an intrinsic transcription factor molecular program and by extrinsic cues (Belecky-Adams et al., 1996; Cepko et al., 1996; Reh and Kljavin, 1989; Watanabe and Raff, 1990). Loss- and gain-of-function studies in model organisms have revealed that the transcription factors Pax6, Rax (also called Rx), Lhx2, Otx2, Sox2, Six6 and Six3 are involved in early eye patterning and retinal developmental processes (Bernier et al., 2000; Carl et al., 2002; Chow et al., 1999; Lagutin et al., 2003; Loosli et al., 1999; Marquardt et al., 2001; Mathers et al., 1997; Porter et al., 1997; Taranova et al., 2006). Later on, specific sets of transcription factors define retinal cell type identity, including photoreceptors (Swaroop et al., 2010). Photoreceptor progenitor and precursor cells express Otx2 and Crx, and conditional deletion of *Otx2* in the developing mouse retina impairs photoreceptors fate (Nishida et al., 2003). In turn, Crx is required for terminal differentiation and maintenance of photoreceptors and is mutated in human retinal degenerative diseases (Chen et al., 1997; Freund et al., 1997, 1998; Furukawa et al., 1997, 1999; Swaroop et al., 1999). Photoreceptors exist as two cell types – rods and cones. Rods are involved in low-intensity night vision and cones are involved in high-intensity color vision. During development, photoreceptors follow an S-cone (cones containing S-opsin) default pathway, which is determined by cone-rod homeobox (Crx) and thryroid receptor β2 (Thrβ2, encoded by *Thrb*); Crx induces expression of *Opn1sw* (encoding S-opsin) by default, whereas Thrβ2 suppresses it and induces expression of *Opn1mw* (encoding M-opsin) (Ng et al., 2001; Yanagi et al., 2002). In turn, expression of neural retina leucine zipper (*Nrl*), RAR-related orphan receptor β (*Rorb*) and *Notch1* inhibit cone formation, whereas both Nrl and RORβ promote rod genesis at the expense of cones (Jadhav et al., 2006; Jia et al., 2009; Mears et al., 2001; Yaron et al., 2006).

Although the transcription factor dynamics controlling retinal development has been well described, the role of chromatin remodeling factors in retinal biology has been poorly explored (Hennig et al., 2013). Polycomb group proteins form large multimeric complexes that silence specific target genes by modifying chromatin organization (Valk-Lingbeek et al., 2004). The Polycomb group protein Bmi1 is a component of the polycomb repressive complex 1 (PRC1), which promotes chromatin compaction and gene repression through its mono-ubiquitin ligase activity on histone H2A at lysine 119 (Buchwald et al., 2006; Cao et al., 2005; Li et al., 2006). *Bmi1*$^{-/-}$ mice show axial skeleton defects, reduced post natal growth and lifespan, and progressive cerebellar degeneration (Jacobs et al., 1999; van der Lugt et al., 1994). Most Bmi1 functions in normal development and stem cell maintenance have been attributed to transcriptional repression of the *Cdkn2a* (also called the *INK4a/ARF*) locus, encoding p16^{INK4a} and p19^{ARF} (Sherr, 2001; Sharpless et al., 2004; Valk-Lingbeek et al., 2004). p16^{INK4a} is a cyclin-dependent kinase inhibitor that blocks the activity of Cdk4/6 by preventing its association with cyclin D, which results in Rb hypophosphorylation and cell cycle arrest or senescence. p19^{ARF} binds and inhibits the activity of the E3-ubiquitin ligase mouse double minute 2 (Mdm2), which prevents targeting of p53 for proteasomal degradation (Sherr, 2001; Sharpless et al., 2004). More recently, activation of the DNA damage response protein checkpoint kinase 2 (Chk2) was found to contribute to several pathologies found in *Bmi1*$^{-/-}$ mice (Liu et al., 2009). *Bmi1*$^{-/-}$ mice also develop a progeroid phenotype in the

[1]Stem Cell and Developmental Biology Laboratory, Hôpital Maisonneuve-Rosemont, 5415 Boul. l'Assomption, Montréal, Canada H1T 2M4. [2]Department of Cognitive Brain Science, National Institute for Longevity Sciences, National Center for Geriatrics and Gerontology, 36-3 Gengo, Morioka, Obu, Aichi 474-8522, Japan. [3]Department of Ophthalmology, Otorhinolaryngology and Cervico-Facial Surgery, Faculty of Medicine, Université Laval, Laval, Canada G1V 0A6. [4]Department of Neurosciences, Université de Montréal, Montréal, Canada H3T 1J4. [5]Department of Ophthalmology, Université de Montréal, Montréal, Canada H3T 1J4.
*These authors contributed equally to this work

‡Author for correspondence (gbernier.hmr@ssss.gouv.qc.ca)

Received 13 April 2015; Accepted 1 March 2016

Development (2016) 143, 1571-1584 doi:10.1242/dev.125351

CNS characterized by lens cataracts, cortical neurons apoptosis, p53 activation and accumulation of oxidative damage (Chatoo et al., 2009). In the developing retina, it was found that Bmi1 is not required for the proliferation of the main RPC population, but for proliferation and post natal maintenance of most immature RPCs located at the retinal ciliary margin. *Bmi1* overexpression in RPCs with short-term proliferating activity induces chromatin remodeling and conversion into long-term RPCs with stem cell characteristics (Chatoo et al., 2010). *Bmi1* thus distinguishes most immature progenitor/stem cells from the main RPC population during retinal development (Chatoo et al., 2010). Notably, *Bmi1* is also expressed in differentiated retinal neurons, including photoreceptors, raising the possibility that it might be important for their post natal development or maintenance (Chatoo et al., 2009, 2010).

We report here that whereas retinal cell type genesis occurs relatively normally in *Bmi1*$^{-/-}$ mice, cone bipolar neurons and cone photoreceptors rapidly degenerate during post natal eye development through necroptosis. In the human retina, *BMI1* was preferentially expressed in cones and *BMI1* inactivation in human embryonic stem cells impaired *CRX* expression and cone terminal differentiation. The cellular phenotype was also associated with chromatin compaction anomalies, activation of tandem repeats and induction of *p53*, as found in *Bmi1*$^{-/-}$ mice. These findings revealed new and partly conserved biological functions for Bmi1 during cone photoreceptor development between mouse and man.

RESULTS

To investigate a possible function of Bmi1 in retinal neurons differentiation, we performed immunohistochemistry (IHC) and immunofluorescence (IF) analyses on retinas from wild-type (WT) and *Bmi1*$^{-/-}$ littermates at post natal day 30 (P30). This revealed that retinal organization and cell type genesis were possibly perturbed in *Bmi1*$^{-/-}$ mice (Figs 1,2 and Figs S1,S2). The distribution of syntaxin, which labels amacrine neurons, was abnormal, resulting in its focal accumulation in the inner nuclear layer (Fig. 1A, Fig. S2). The distribution of S-opsin, which labels the outer segment of cones, was also perturbed (Fig. 1A, Fig. S1). By contrast, the overall intensity and distribution of rhodopsin, which labels the rod outer segment, was apparently normal. The total number of nuclei in the outer nuclear layer, which is a measure of the total number of photoreceptors (mostly rods), was also comparable between the two genotypes (Fig. 1A,B, Fig. S2). We quantified the number of all major retinal cell types using specific antibodies (Haverkamp and Wässle, 2000; Marquardt, 2003; Marquardt et al., 2001). We found that the number of ganglion, amacrine and horizontal cells was comparable between both genotypes (Fig. 1G,H, Fig. S2). However, the number of S-cones as well as rod bipolar (PKCα^{+}/Chx10^{+}) and cone bipolar (Chx10^{+}/PKCα^{-}) neurons was highly reduced in *Bmi1* mutants (Fig. 1B, Fig. 2A-C). Amongst the two subcategories of cone bipolar neurons, T2 ON and T8 OFF neurons (labeled by recoverin) were also nearly absent (Fig. 2G, Fig. S2).

To test whether the observed reduction in cone numbers in *Bmi1*$^{-/-}$ mice was the result of a developmental defect or secondary to degeneration, we compared P12 and P30 mice. On retinal flat mounts, we found that S-opsin^{+} and PNA^{+} cone cells were equally abundant in the ventro-nasal region of WT and *Bmi1*$^{-/-}$ mice at P12, but not at P30, thus suggesting cone degeneration (Fig. 1C-F, Fig. S1). The apparent over-representation of S-opsin in the ventro-nasal region, in contrast to the dorso-temporal region, is due to the dual nature of mouse cone photoreceptors, which frequently express both S-opsin and M-ospin (Fig. S4A) (Ortín-Martínez et al., 2014). These analyses also highlighted a significant reduction in cone numbers in *Bmi1*$^{+/-}$ mice, revealing a gene-dosage effect (Fig. 1C,D). We performed labeling with an antibody against activated caspase-3 to test for apoptosis. Positive cells were, however, not observed in either WT or *Bmi1*$^{-/-}$ retinas at P30, suggesting that apoptosis is not the main mechanism of retinal cell death in *Bmi1*$^{-/-}$ mice (data not shown). This is consistent with our previous findings, where differences in the number of caspase-3^{+} cells were not observed between WT and *Bmi1*$^{-/-}$ retinas at P6 (Chatoo et al., 2010). To test for an altered retinal differentiation program, we performed quantitative RT-PCR (qPCR) on retinal extracts. Although expression of the Bmi1-repressed locus *Cdkn2a* (encoding the *p16*^{INK4a} and *p19*^{ARF} transcripts) was increased in *Bmi1*$^{-/-}$ retinas (Fig. 2I) (Molofsky et al., 2005), no difference in the expression level of several retinal determination genes was observed between the two genotypes.

We also investigated *Bmi1* expression in the WT mouse retina at P30. Although *Bmi1* was expressed in nearly all retinal cell types, it was most abundant in Chx10^{+} bipolar neurons (Fig. S5). In the outer nuclear layer, it was predominant in cones, not in rods, as further shown using the cone-only retina of *Nrl* null mice (Fig. S3A) (Mears et al., 2001). To test whether the cone degeneration phenotype was cell autonomous, we performed dissociated retinal cultures from WT and *Bmi1*$^{-/-}$ mice at P1. Although the percentage of S-cones was comparable after 4 days *in vitro* (DIV) between both genotypes, as revealed using Bmi1 and S-opsin immunolabeling (Fig. S3B,C), it was highly reduced in *Bmi1*$^{-/-}$ cultures after 8 DIV and 12 DIV, suggesting cell-autonomous degeneration of cones (Fig. S3C). Taken together, these observations revealed predominant expression of Bmi1 in retinal neurons that degenerate in *Bmi1*$^{-/-}$ mice.

Cone function is severely perturbed in *Bmi1*$^{-/-}$ mice

To test the visual function, we performed electroretinogram (ERG) recordings of mouse retinal activity at P30. Examination of the ERG traces for the cone system revealed that cone activity was severely affected in *Bmi1*$^{-/-}$ mice when compared with WT (Fig. 3A-C). Notably, we observed a dose-dependent effect as cone activity was also perturbed in *Bmi1*$^{+/-}$ mice. When tracing the luminance response functions (LRFs), significant genotype-dependent differences were observed between WT, *Bmi1*$^{+/-}$ and *Bmi1*$^{-/-}$ mice for all four ERG parameters in the photopic condition, namely the a-wave amplitude ($F_{2,7}$=6.195, P=0.028; Fig. 3B, Table S1), the b-wave amplitude ($F_{2,7}$=18.970, P=0.001; Fig. 3C, Table S1), the a-wave implicit time ($F_{2,7}$=5.127, P=0.043; Table S1) and the b-wave implicit time ($F_{2,7}$=20.852, P=0.001; Table S1). These results suggest that cone photoreceptor activity, as measured with the a-wave, and the cone bipolar cell activity, as measured with the b-wave, were severely altered in *Bmi1*$^{-/-}$ mice. Also, the rod system was affected in *Bmi1*$^{-/-}$ mice, as observed by the lower b-wave amplitude when compared with WT mice, but the difference was not significant for *Bmi1*$^{+/-}$ mice (Fig. 3D-F, Table S1). When analyzing the LRFs, only the b-wave amplitude was found to be significantly different between WT and *Bmi1*$^{-/-}$ mice ($F_{2,7}$=10.440, P=0.008; Fig. 3F, Table S1), suggesting that rod bipolar cell activity was affected in *Bmi1*$^{-/-}$ mice, but at a lower level than in photopic conditions (Fig. 3C).

Loss of heterochromatin and necroptosis in *Bmi1*$^{-/-}$ cones

We next analyzed the retina of *Bmi1*$^{-/-}$ mice by transmission electron microscopy (TEM) to search for ultrastructural anomalies at P30. In WT mice, the electron-dense heterochromatin of rods was pre-eminent, located in the center of the nucleus and

Development (2016) 143, 1571-1584 doi:10.1242/dev.125351

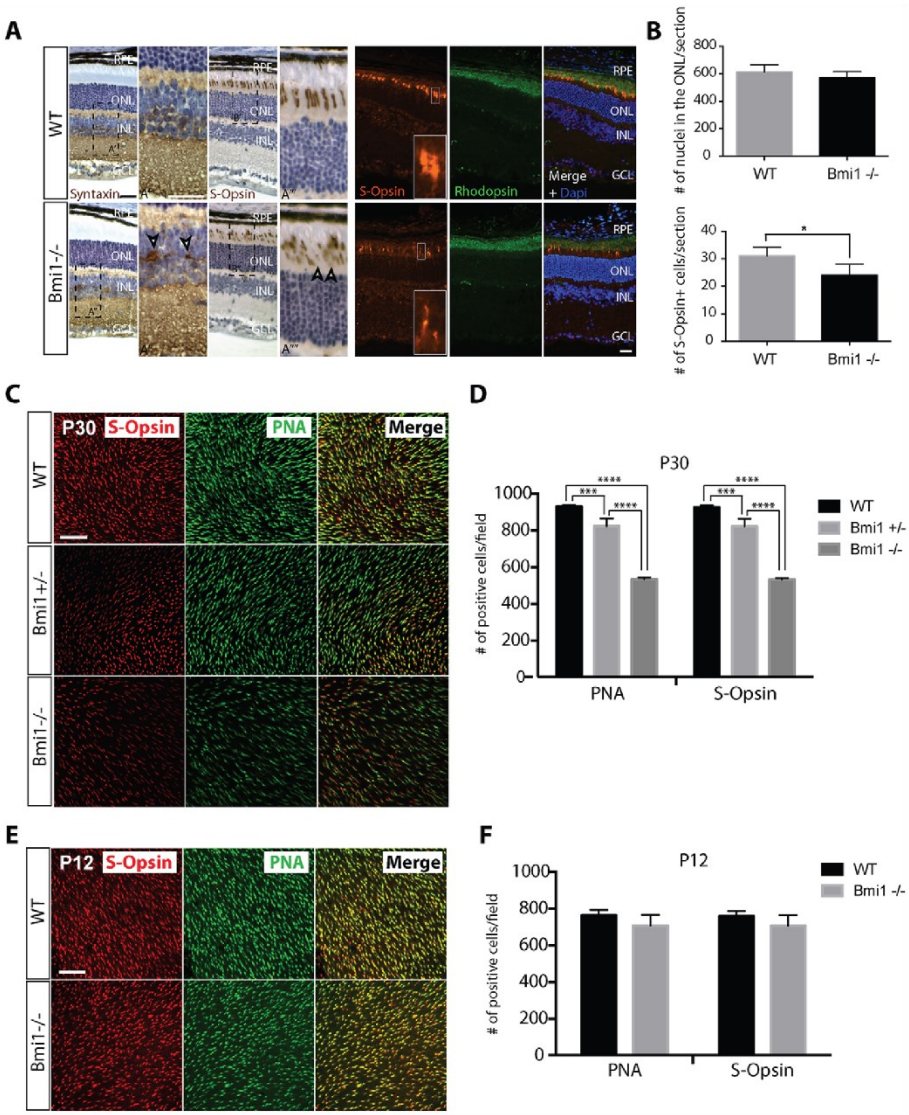


Fig. 1. S-cones degenerate during post natal eye development in *Bmi1*$^{-/-}$ mice. (A) Analyses of WT and *Bmi1*$^{-/-}$ retinas at P30. Abnormal distribution of syntaxin (arrowheads in A″) and cone photoreceptor outer segments breaks in *Bmi1*$^{-/-}$ mice (arrowheads in A‴) as visualized by IHC. Cone photoreceptor outer segments break but rod photoreceptors appear normal in *Bmi1*$^{-/-}$ mice, as visualized by IF. (B) Quantification of total number of photoreceptors (top) and S-cones (bottom). (C,E) Representative microscopy images from retinal flat-mount of WT, *Bmi1*$^{+/-}$ and *Bmi1*$^{-/-}$ mice at P30 (C) and WT and *Bmi1*$^{-/-}$ at P12 (E). Images were taken in the ventro-nasal part of the retina. (D,F) Quantification of S-cone photoreceptors (S-opsin^{+}) and total cone photoreceptors (PNA^{+}) at P30 (D) and P12 (F). RPE, retinal pigment epithelium; ONL, outer nuclear layer; INL, inner nuclear layer; GCL, ganglion cell layer. Scale bars: 40 µm. All values are means±s.e.m. *P≤0.05; ***P≤0.001; ****P≤0.0001; Student's *t*-test (B,F), two-way ANOVA (D).

generally organized as a single large chromocenter (Fig. 4A, gray arrows) (Solovei et al., 2009). By contrast, cone nuclei had 1-3 chromocenters that occupied a much-reduced nuclear surface (Fig. 4A, white arrows). Cone cell bodies were also located close to the junction with the inner segment, in contrast to rod cell bodies, which were evenly distributed in the outer nuclear layer (Fig. 4A, Fig. S3A). In *Bmi1*$^{-/-}$ mice, the nucleus of rods appeared normal but the chromocenter was slightly reduced in size

Development (2016) 143, 1571-1584 doi:10.1242/dev.125351

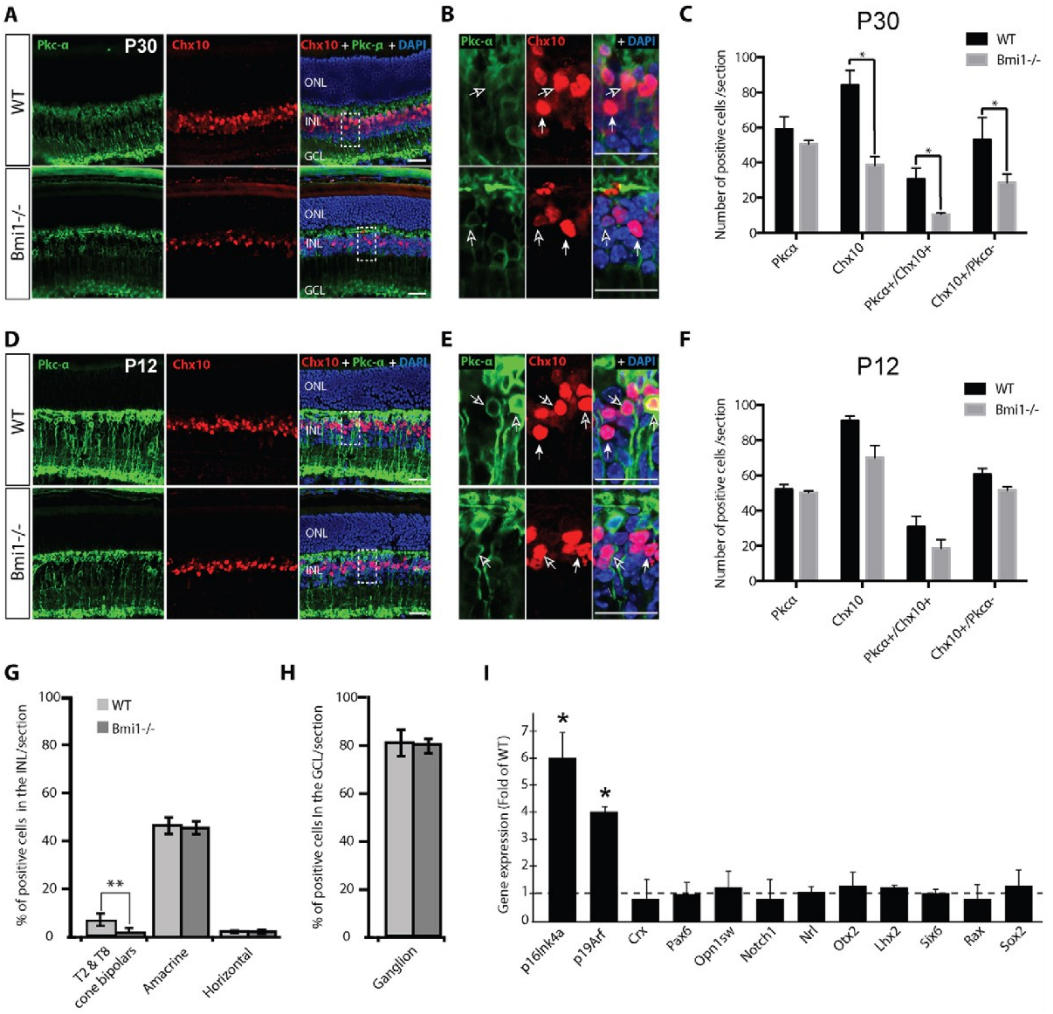


Fig. 2. Bipolar neurons degenerate during post natal eye development in *Bmi1*$^{-/-}$ mice. (A,D) Representative images of WT and *Bmi1*$^{-/-}$ retinas at P30 (A) and P12 (D) labeled with antibodies against PKCα and Chx10. (B,E) Cropped images indicated by the respective dashed rectangles indicated in A and D. White filled arrows: cone bipolar cells (Chx10^{+}/PKCα^{-}). White-edged arrows: rod bipolar cells (Chx10^{+}/PKCα^{+}). Note the decrease in Chx10^{+} cells in *Bmi1*$^{-/-}$ retinas at P30. (C,F) Quantification of data acquired in A and D. (G) Quantification of T2-OFF and T8-ON cone bipolar cells (Recoverin^{+}), amacrine cells (Pax6^{+}) and horizontal cells (Calbindin^{+}) in the INL. (H) Quantification of ganglion cells (Pax6^{+}) in the GCL. (I) Gene expression analysis of WT and *Bmi1*$^{-/-}$ retinas at P30, where *p16*^{INK4a} and *p19*^{ARF} were used as positive controls. RPE, retinal pigment epithelium; ONL, outer nuclear layer; INL, inner nuclear layer; GCL, ganglion cell layer. Scale bars: 40 µm. All values are means±s.e.m. *P≤0.05; **P≤0.01; Student's *t*-test.

and showed less condensation (Fig. 4B). The cell body, inner segment and outer segment of rods were apparently normal. By contrast, the cell bodies and nuclei of cones were swollen and highly degenerative (Fig. 4A, white arrows). Accumulation of swollen mitochondria and fibrous material in the cell body of cones was also observed, suggesting necrotic cell death (also known as necroptosis) (Fig. 4A, white and black arrows in the high-magnification image) (Linkermann and Green, 2014).

To further investigate the observed chromatin phenotype, we analyzed mice at P15, when chromatin condensation of immature rods is not yet completed. Marked differences in rod heterochromatin condensation could be observed between WT and *Bmi1*$^{-/-}$ mice by transmission electron microscopy (TEM) (Fig. 4B). Using retinal sections at P25 and antibodies against 'open' chromatin (H3K9ac), facultative heterochromatin (H3K27me3) and constitutive heterochromatin (H3K9me3), we found that the chromatin in the outer nuclear layer of *Bmi1*$^{-/-}$ mouse retinas was less condensed

Development (2016) 143, 1571-1584 doi:10.1242/dev.125351

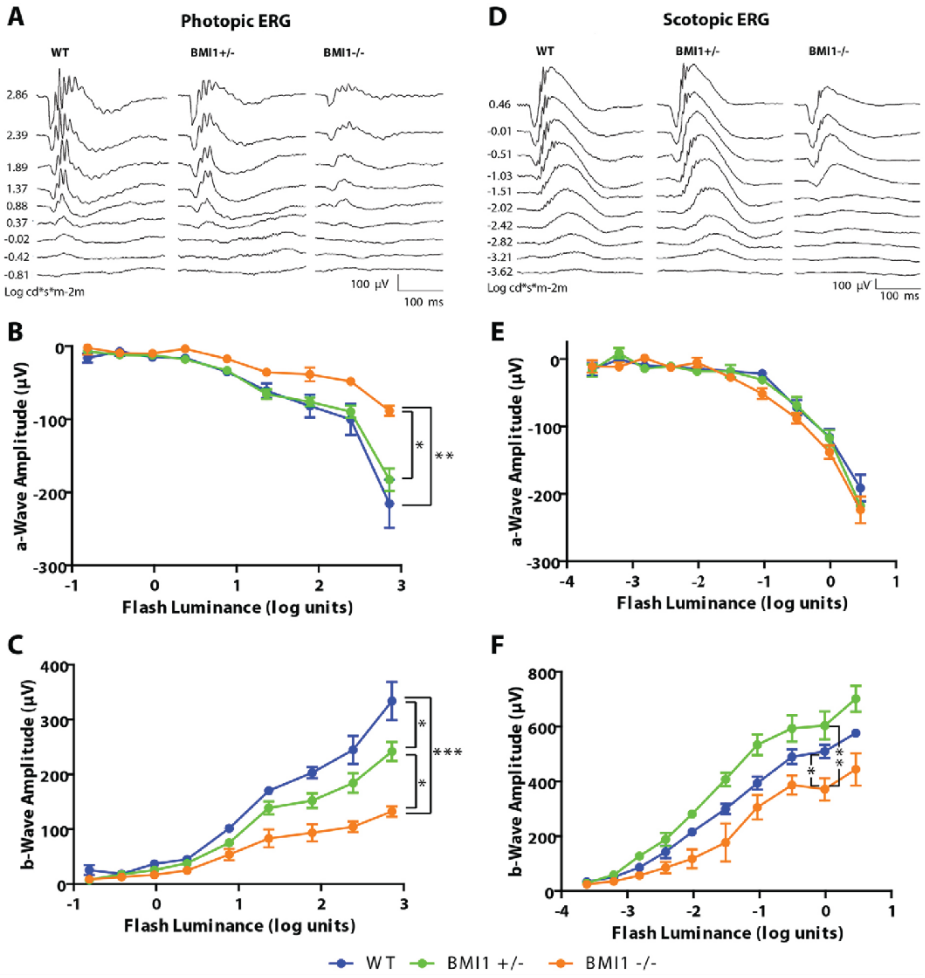


Fig. 3. Cone and rod function is severely perturbed in *Bmi1*$^{-/-}$ mice. ERG recorded in photopic (A-C) and scotopic (D-F) conditions. (A,D) ERG waveforms for WT, *Bmi1*$^{+/-}$ and *Bmi1*$^{-/-}$ mice. Luminance response function curve of a-wave (B,E) and b-wave (C,F) representing the dynamic electrical response from the cone (B,C) and rod (E,F) systems. Note the significant decrease in the retinal responses of *Bmi1*$^{+/-}$ mice and *Bmi1*$^{-/-}$ mice in photopic conditions for both the a-wave and b-wave amplitude at the V_{max} compared with WT mice and the decrease of the b-wave amplitude at the V_{max} in scotopic conditions in *Bmi1*$^{-/-}$ mice compared with WT and *Bmi1*$^{+/-}$ mice. All values are means±s.e.m. *P≤0.05; **P≤0.01; ***P≤0.001; one-way ANOVA.

when compared with that of WT mice (Fig. 4C) (Fodor et al., 2010). The intergenic and pericentromeric constitutive heterochromatin contains numerous repetitive DNA sequences of retroviral origin that can be transcribed but are generally silenced by heterochromatin formation (Fodor et al., 2010; Karimi et al., 2011). We thus compared the expression level of repetitive DNA sequences (Iap1, Line, Sine, major satellite repeats and minor satellite repeats) between retinas from WT and *Bmi1*$^{-/-}$ mice by qPCR using RNA extracts first treated with DNaseI, because these transcripts are intron-less. We found increased expression of all tested repetitive sequences in *Bmi1*$^{-/-}$ retinas, consistent with the reduced heterochromatin compaction phenotype (Fig. 4D).

To investigate the mechanism of cell death, we measured the expression of genes known to mediate necroptosis (Murakami et al., 2012; Vandenabeele et al., 2010; Viringipurampeer et al., 2014). We found that the mRNA level of receptor-interacting protein kinase 3 (*Ripk3*) was significantly increased in *Bmi1*$^{-/-}$ retinas when compared with WT (Fig. 5A). Rip3 protein accumulation in *Bmi1*$^{-/-}$ and *Bmi1*$^{+/-}$ retinas was confirmed by western blot analysis (Fig. 5B). In these preparations, we noticed that Rip3 accumulation could be readily observed upon Ponceau Red staining of the nitrocellulose membrane. Furthermore, protein accumulation was present at ~20 kDa in *Bmi1*$^{-/-}$ retinal extracts, suggesting 'programmed' proteolytic cleavage (Fig. 5B). To

Development (2016) 143, 1571-1584 doi:10.1242/dev.125351

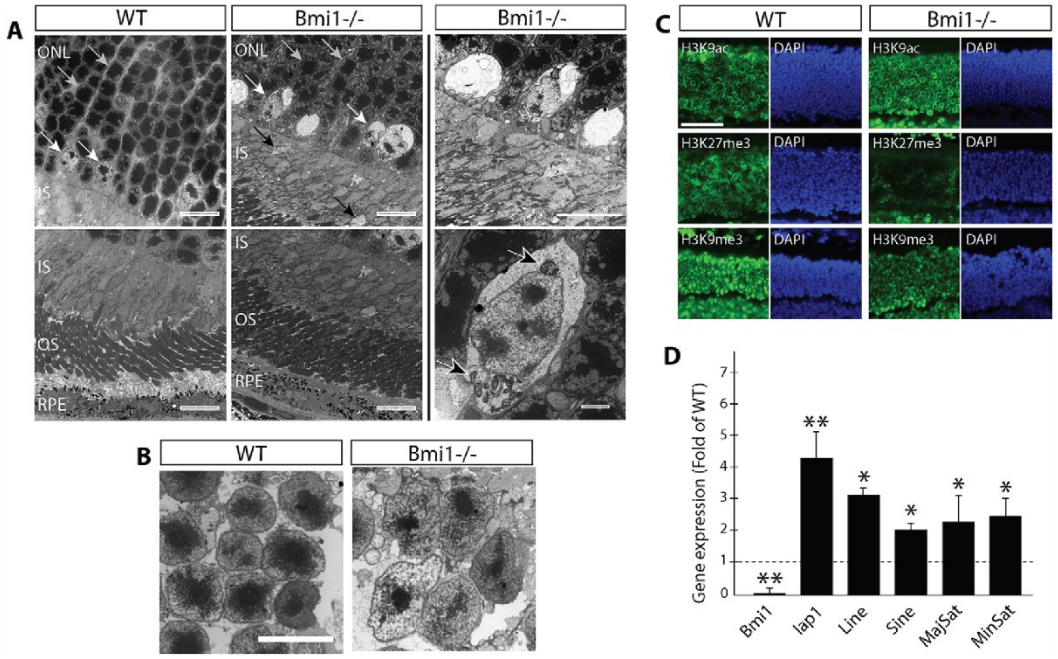


Fig. 4. Cone necrosis, loss of heterochromatin and activation of tandem repeats in *Bmi1*$^{-/-}$ mouse retinas. (A) TEM on WT and *Bmi1*$^{-/-}$ retinas at P30. White arrows, cone photoreceptors; gray arrows, rod photoreceptors characterized by electron-dense heterochromatin organized as a single chromocenter; black arrows, degenerative cone cell bodies; white-edged black arrows, swollen mitochondria in cones. (B) TEM analysis of WT and *Bmi1*$^{-/-}$ rod photoreceptors at P15. (C) Analysis of histone modifications in WT and *Bmi1*$^{-/-}$ retinas by IF on sections at P25. (D) Quantitative PCR analysis of expression of repetitive DNA sequences in WT and *Bmi1*$^{-/-}$ retinas at P30. *Bmi1* is used as a negative control and values are expressed as fold of WT value. RPE, retinal pigment epithelium; ONL, outer nuclear layer; OS, outer segment; IS, inner segment. Scale bars: 10 µm (A, left, center and top-right images); 2 µm (A, bottom-right image); 5 µm in B; 25 µm in C. All values are means±s.e.m. *P≤0.05; **P≤0.01; Student's *t*-test.

identify which retinal cell types were most affected, we performed immunofluorescence (IF) analyses on sections. In both P12 and P30 *Bmi1*$^{-/-}$ retinas, we observed Rip3 immunolabeling in cones, but not in rods, and in neurons located in the inner nuclear layer, thus possibly corresponding to bipolar neurons (Fig. 5C-G). Rip3 immunolabeling was observed in the cell body and outer segment of cones at P12 (Fig. 5E,F), but was predominant in the outer segment of cones at P30 (Fig. 5C,D). Morphological anomalies of the cone outer segment were also clearly visible in *Bmi1*$^{-/-}$ retinas at P30 (Fig. 5D). Taken together, these results reveal that necroptosis is the main mechanism of cone photoreceptors cell death in *Bmi1*$^{-/-}$ mice.

Chk2 or *p53* genetic deficiency can partially improve the *Bmi1*$^{-/-}$ retinal phenotype

To investigate additional molecular mechanisms leading to cone degeneration, we generated double null mutants for the *Bmi1* and *p16*^{INK4a}/*p19*^{ARF}, *p19*^{ARF} or *Chk2* alleles. As previously reported, the *p16*^{INK4a}/*p19*^{ARF} or *p19*^{ARF} mutant alleles did not rescue size or lifespan defects in *Bmi1* null mice (Molofsky et al., 2005). By contrast, insertion of the *Chk2* mutant allele improved size and lifespan in *Bmi1* null mice, although *Bmi1*$^{-/-}$/*Chk2*$^{-/-}$ mice remained smaller than normal (Fig. S6). Interestingly, the number of cones in *Bmi1* null retinas at P30 was significantly improved by the *Chk2* mutant allele, but not by the *p16*^{INK4a}/*p19*^{ARF}or *p19*^{ARF} mutant alleles (Fig. 6A,B). However, morphology of the cone outer segment remained highly abnormal in *Bmi1*$^{-/-}$/*Chk2*$^{-/-}$ mice (Fig. 6A). Hence, analysis of *Bmi1*$^{-/-}$/*Chk2*$^{-/-}$ mice at P150 revealed severe depletion of S-cones, suggesting that deletion of *Chk2* only provided a transitory rescue (Fig. 6C,D).

In cortical neurons of *Bmi1* null mice, stabilization of p53 leads to accumulation of reactive oxygen species (ROS) through repression of the *Nqo1*, *Gsta1* and *Sesn2* antioxidant genes (Chatoo et al., 2011, 2009). p53 also promotes neuronal cell death through activation of *Apaf1*, *Fas* and *Lpo* (Fortin et al., 2001) and Bmi1 was proposed to block ROS accumulation in blood cells and thymus through direct transcriptional repression of the pro-oxidant genes *Cyp24a1* and *Duox2* (Liu et al., 2009). By qPCR analysis on retinal extracts, we found that the expression of antioxidant genes in *Bmi1*$^{-/-}$ mice was unchanged or increased (*Nqo1*), whereas that of *Cyp24a1* and *Duox2* was increased (Fig. 7A). Notably, the expression of *Apaf1* and *Fas* was increased in *Bmi1*$^{-/-}$ retinas, consistent with the observed accumulation of p53 by western blot (Fig. 7B). Among 200 offspring, we obtained expected numbers of WT, *Bmi1*$^{+/+}$/*p53*$^{+/-}$, *Bmi1*$^{+/+}$/*p53*$^{-/-}$, *Bmi1*$^{+/-}$/*p53*$^{+/-}$, *Bmi1*$^{+/-}$/*p53*$^{+/+}$, *Bmi1*$^{-/-}$/*p53*$^{+/+}$ and *Bmi1*$^{-/-}$/*p53*$^{+/-}$ mice, but a single *Bmi1*$^{-/-}$/*p53*$^{-/-}$ mouse (expected n=12), for which size and viability did not improve. To evaluate the contribution of *p53* to the cone degeneration phenotype of *Bmi1*$^{-/-}$ mice, we performed double staining with S-opsin and peanut agglutinin (PNA), which labels the outer segments of all

Development (2016) 143, 1571-1584 doi:10.1242/dev.125351

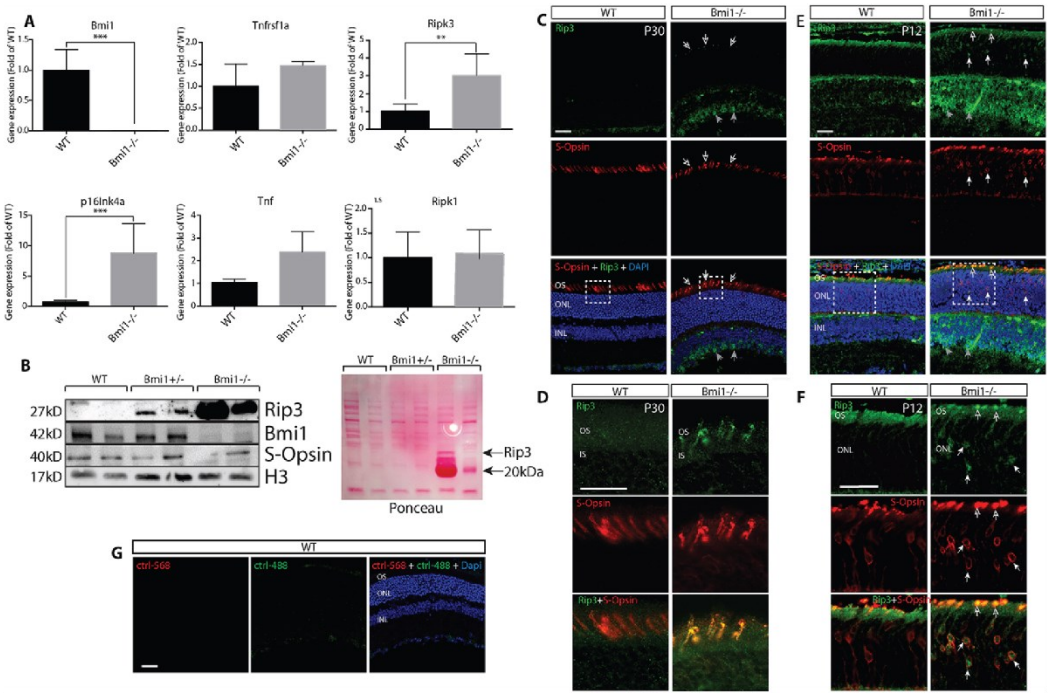


Fig. 5. Cone degeneration in *Bmi1*$^{-/-}$ mice operates through necroptosis. (A) RT-qPCR analysis of *Bmi1*$^{-/-}$ and WT retinal extracts at P30; *Bmi1* and *P16Ink4a* were used as internal controls. Values are expressed as fold of WT. (B) Western blot analyses of retinal extracts from WT, *Bmi1*$^{+/-}$ and *Bmi1*$^{-/-}$ mice. Note the gene dosage-dependent increase in Rip3 expression and the decrease in S-opsin expression in *Bmi1*$^{-/-}$ mice. Ponceau-red staining of the membrane is shown on the right. Histone H3 was used for protein normalization. (C-F) IF analyses on WT and *Bmi1*$^{-/-}$ retinas at P30 (C,D) and P12 (E,F). White-edged arrows, accumulation of Rip3 in S-cone outer segments; gray arrows, Rip3^{+} cells in the INL; white arrows, S-opsin^{+} cells with nuclear expression of Rip3 in *Bmi1*$^{-/-}$ mice at P12 not present at P30. (G) Control with secondary antibody only. INL, inner nuclear layer; ONL, outer nuclear layer; OS, outer segment; IS, inner segment. Scale bars: 40 µm. All values are means±s.e.m. $^{**}P \leq 0.01$; $^{***}P \leq 0.001$; Student's *t*-test.

cones. By confocal microscopy, we reconstructed retinal sections in 3D to calculate the number of cones and evaluate their morphology. Interestingly, while the number of S-opsin^{+}/PNA^{+} cells was reduced in *Bmi1*$^{-/-}$/*p53*$^{+/-}$ retinas when compared with WT, it was improved in the unique *Bmi1*$^{-/-}$/*p53*$^{-/-}$ retina sample (Fig. 7C,D). However, S-opsin labeling in the *Bmi1*$^{-/-}$/*p53*$^{-/-}$ retina remained fragmented and disorganized, suggesting improved survival, but not morphology, of S-cones (Fig. 7C). We also analyzed samples using the MitoSoxRed reagent, which reacts with mitochondrial ROS on unfixed tissue. In *Bmi1*$^{-/-}$/*p53*$^{+/-}$ retinal sections, we observed robust fluorescence when compared with WT, suggesting increased mitochondrial ROS. Notably, fluorescence was highly reduced in the *Bmi1*$^{-/-}$/*p53*$^{-/-}$ retina but was not completely restored to WT levels (Fig. 7E). These results revealed the partial contribution of *Chk2* and *p53* to the cone degeneration phenotype of *Bmi1*$^{-/-}$ mice.

BMI1 is enriched at heterochromatic foci in human cones

We previously reported *BMI1* expression in the human retina and its downregulation during aging (Abdouh et al., 2012). Here, we investigated localization of BMI1 in human photoreceptors. Using confocal IF analyses on adult human retinas, we observed punctuate BMI1 immunolabeling in the nucleus of photoreceptors (Fig. 8A-C). In some cells that co-labeled with S-opsin or M-opsin, BMI1 immunolabeling was even more robust and present as multiple foci per nucleus (Fig. 8A). To characterize the BMI1 immunolocalization pattern on the chromatin of cones, we performed dual immunolabeling with antibodies directed against distinct histone modifications. Although BMI1 did not colocalize with H3K9ac, it colocalized with H3K27me3 and H3K9me3, suggesting distribution at both facultative and constitutive heterochromatin in human cones (Fig. 8B).

BMI1 is required for human cone differentiation and chromatin integrity

To investigate the function of BMI1 during human cone development, we used a protocol allowing the differentiation of ~70% of human embryonic stem (hES) cells into cones (Zhou et al., 2015). This method results in the generation of immature S-cones expressing CRX (*CRX*), cone arrestin (*ARR3*) and S-opsin (*OPN1SW*) within 21 days (Fig. 9A,D,E). Importantly, *in vitro* generated S-cones expressed BMI1, which also colocalized with H3K9me3 (Fig. 9A,B). We infected hES cells with a lentivirus expressing a small hairpin RNA against *BMI1* (shBMI1) or scramble sequence (shScramble) and performed hygromycin selection for 10 days (Abdouh et al., 2009). The hES cells were

Development (2016) 143, 1571-1584 doi:10.1242/dev.125351

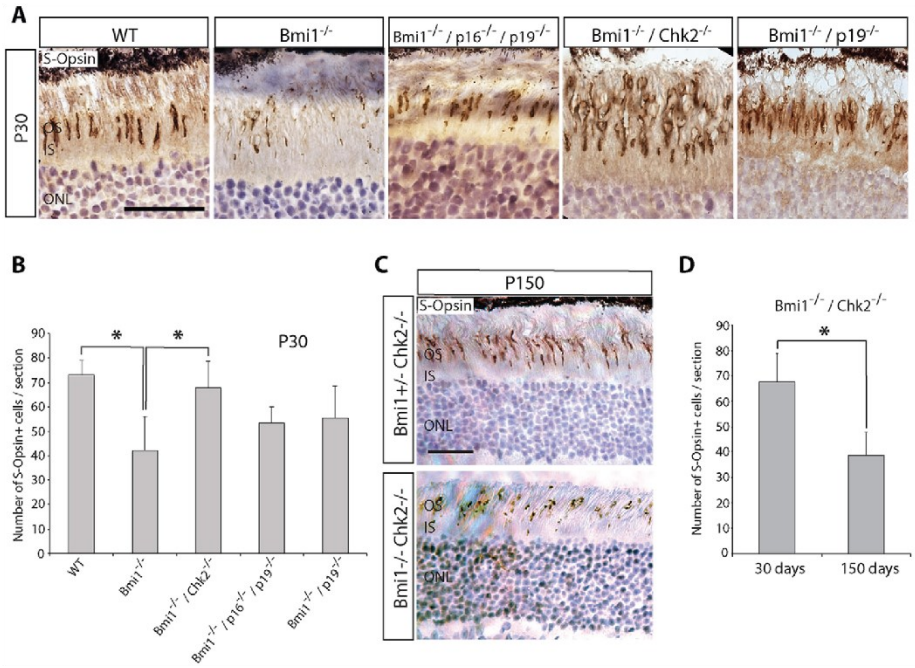


Fig. 6. *Chk2* genetic deficiency improves the *Bmi1*$^{-/-}$ cone degeneration phenotype. (A) Immunohistochemistry analyses of WT, *Bmi1*$^{-/-}$ and double mutant retinas at P30. (B) Quantification of the number of S-opsin^{+} photoreceptors per section for the five genotypes at P30 (n=4 mice for each genotype). (C) IHC analyses of retinas at P150. (D) Quantification of the number of S-opsin^{+} cells per field at P30 and P150. Scale bars: 40 µm. Quantifications are made from ×25 images. All values are means±s.e.m. *P≤0.05; Student's t-test.

then differentiated into S-cones for 21 days. Upon western blot analysis, BMI1 expression was hardly visible in naive hES cells, in contrast to *in vitro* differentiated cones (Fig. 9C). Histone H2Aub is the target of BMI1 and RING1a/b biochemical activity (Buchwald et al., 2006; Cao et al., 2005; Li et al., 2006; Wang et al., 2004). Consistently, histone H2Aub was reduced in differentiated cones upon knockdown of BMI1 (Fig. 9C). To confirm cone differentiation, we analyzed cells for expression of S-opsin and CRX (Chen et al., 1997; Freund et al., 1997, 1998; Furukawa et al., 1997, 1999). In control cells, S-opsin and CRX expression was observed in more than 70% of cells, as visualized by IF and western blot (Fig. 9A and D). By contrast, S-opsin and CRX expression was highly reduced in differentiated cones upon BMI1 knockdown (Fig. 9A,D). When using a hypomorphic construct against BMI1 that reduces BMI1 levels by ∼50% (shBMI1 50%) (Abdouh et al., 2009), CRX expression in differentiated cones was reduced by ∼25% when compared with levels in control cells, suggesting a modest gene-dosage effect (Fig. 9D). To consider the role of p53 activation in the context of *Bmi1* deficiency, we investigated p53 expression (Chatoo et al., 2009). Using an antibody against p53 that also recognizes other family members, i.e. p73 and p63, we observed predominant expression of p73 in shScramble and shBMI1 50% cones (Fig. 9D) (Jacobs et al., 2006). By contrast, p73 was lost in shBMI1 95% cones and was associated with induction of *p63* and *p53* (Fig. 9D). It was previously proposed that one main function of Polycomb group proteins in mouse ES cells is to prevent differentiation through repression of lineage-specific homeobox genes such as PAX, SOX and LHX families (Boyer et al., 2006). To further characterize the cone differentiation phenotype, we analyzed the hES cells by qPCR. We found that while the expression of cone-specific genes (*ARR3* and *OPN1SW*) and retinal homeobox genes (*RAX* and *SIX6*) was reduced by 50-75% in shBMI1 cells, expression of *SOX1* was increased by ∼6-fold when compared with control cells (Fig. 9E). *SOX1* and *SOX2* are enriched in neural stem cells and retinal progenitors but are not expressed in photoreceptors, thus providing a possible explanation for the defective terminal differentiation of cones upon knockdown of BMI1 (Avilion et al., 2003; Ellis et al., 2004; Graham et al., 2003; Taranova et al., 2006; Yan et al., 2005).

To investigate the effect of BMI1 knockdown on the chromatin of human cones, we performed IF analyses. BMI1 knockdown resulted in increased H3K9ac levels, but reduced H3K27me3 and H3K9me3 levels when compared with control cells (Fig. 9F). BMI1 knockdown was also associated with the formation of γH2Ax foci – a histone modification characteristic of DNA double-strand breaks and genomic instability (Fig. 9G) (Chagraoui et al., 2011; Facchino et al., 2010; Ismail et al., 2010; Rogakou et al., 1998). Consistent with the chromatin relaxation phenotype, we observed induction of repeat DNA sequences in BMI1 knockdown cones, with a modest gene-dosage effect when using the hypomorphic shBMI1 construct (Fig. 9H). These experiments revealed that BMI1 is required for terminal differentiation, heterochromatin compaction, silencing of repeat DNA and genomic stability in human cones.

Development (2016) 143, 1571-1584 doi:10.1242/dev.125351

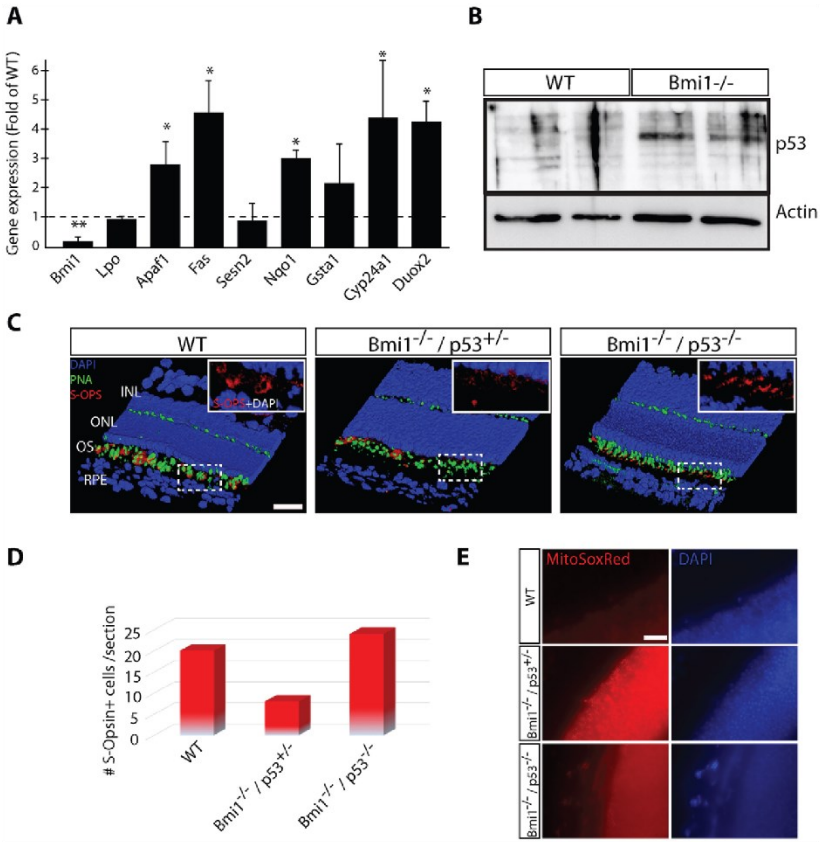


Fig. 7. *p53* genetic deficiency improves the *Bmi1*$^{-/-}$ cone degeneration phenotype. (A) Quantitative RT-PCR analysis of WT and *Bmi1*$^{-/-}$ retinal extracts at P30. *Bmi1* was used as negative control and all values are expressed as fold of WT. (B) Western blot analysis of mouse retinas at P30 showing p53 accumulation in *Bmi1* mutants. (C) 3D reconstruction by confocal microscopy analyses of WT, *Bmi1*$^{-/-}$/*p53*$^{+/-}$ and *Bmi1*$^{-/-}$/*p53*$^{-/-}$ retinas at P30. White box: detail of S-opsin in the outer segment in the dashed box. Note the lower immunoreactivity and the fragmentation of the S-opsin in *Bmi1*$^{-/-}$/*p53*$^{+/-}$ mice and a rescue in the quantity of S-opsin in the *Bmi1*$^{-/-}$/*p53*$^{-/-}$ even if it is still fragmented. (D) Quantification of S-opsin^{+} cells per section in WT, *Bmi1*$^{-/-}$/*p53*$^{+/-}$ and *Bmi1*$^{-/-}$/*p53*$^{-/-}$ mice. (E) IF analyses of WT and *Bmi1*$^{-/-}$ retinas at P30 stained with MitoSoxRed. Note the partial rescue in *Bmi1*$^{-/-}$/*p53*$^{-/-}$ mice when compared with WT. RPE, retinal pigment epithelium; ONL, outer nuclear layer; INL, inner nuclear layer; OS, outer segment. Scale bar: 50 µm in C; 20 µm in E. All values are means±s.e.m. *P≤0.05; **P≤0.01; Student's *t*-test.

DISCUSSION

We showed here that cone photoreceptors and bipolar neurons are normally generated but then undergo rapid degeneration in *Bmi1*$^{-/-}$ mice through Rip3-associated necroptosis. Selective retinal cell degeneration in *Bmi1*$^{-/-}$ mice also correlated with predominant Bmi1 expression in bipolar neurons and cone photoreceptors. Cone number but not morphology in *Bmi1*$^{-/-}$ mice was partially rescued by deletion of either *Chk2* or *p53*, implicating these additional pathways in neurodegeneration. BMI1 was expressed in human cones, where it localized at heterochromatic foci. BMI1 inactivation in hES cells severely perturbed differentiation of cones, in contrast to the situation found in *Bmi1*$^{-/-}$ mice. However, BMI1-deficient human cones also presented common features with the *Bmi1*$^{-/-}$ mouse retinal phenotype, thus revealing partially conserved functions across species.

In previous work, it was proposed that *Bmi1* inactivation could prevent rod and cone photoreceptor degeneration in *Rd1* mice, a model of retinitis pigmentosa (Zencak et al., 2013). *Rd1* mice carry a mutation in *Pde6b*, which is only expressed in rods. Thus, cone degeneration in *Rd1* mice is thought to be secondary to the loss of trophic support provided by rods. As apoptosis of rod photoreceptors in *Rd1* mice is preceded by cell cycle re-entry and activation of cyclin dependent kinases (CDKs), it was suggested that loss of *Bmi1* provides neuroprotection by blocking CDK activation and thus cell cycle re-entry of rods (Marigo, 2007; Sancho-Pelluz et al., 2008; Zencak et al., 2013). Here, we found that *Bmi1* deficiency was associated with the activation of several cell cycle inhibitors in the retina such as p16^{INK4a}, p19^{ARF}, Chk2 and p53, consistent with the general function of Bmi1 in inhibiting the p16^{INK4a}/CDK6/Rb and p19^{ARF}/p53/p21^{Cip1} pathways (Sauvageau and Sauvageau, 2010). Thus, taken in the context of the *Rd1*

Development (2016) 143, 1571-1584 doi:10.1242/dev.125351

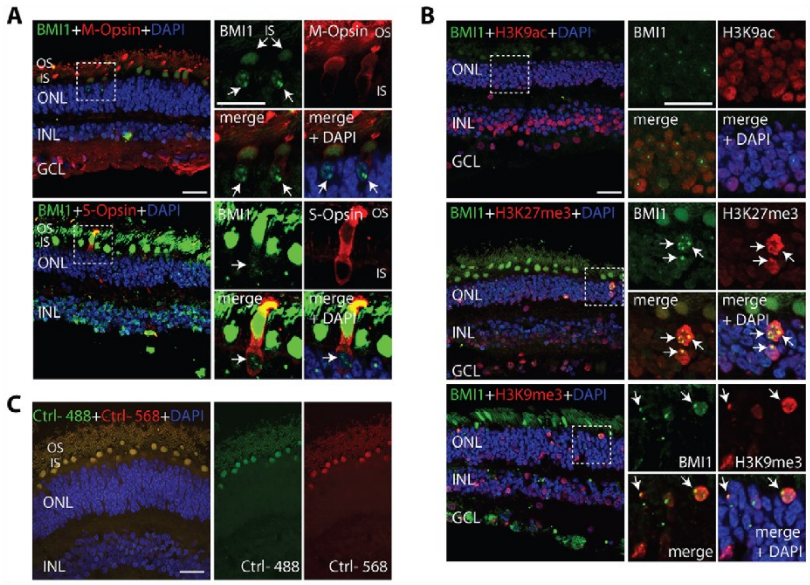


Fig. 8. BMI1 is enriched at heterochromatic nuclear foci in human cones. (A-C) Confocal IF analysis of BMI1 expression in the adult human retina. (A) BMI1 is expressed in the nucleus of M-opsin$^+$ and S-opsin$^+$ cones (arrows). (B) BMI1 is enriched at heterochromatic nuclear foci as shown by colocalization with H3K9me3 and H3K27me3 in cones (arrows). (C) Negative control with only the secondary antibody. Note the strong autofluorescence present in the IS and OS. ONL, outer nuclear layer; INL, inner nuclear layer; IS, inner segment; OS, outer segment; GCL, ganglion cell layer. Scale bars: 20 µm.

mutation, the proposed model is likely to be valid, but only in rods. Also, whether acute Bmi1 inhibition in *Rd1* mice can prevent rod degeneration remains to be demonstrated, as this would be more relevant to a clinical context. Interestingly, the outer segment of cones is abnormal in *Rd1/Bmi1*$^{-/-}$ mice, as shown using S-opsin immunolabeling, suggesting rod-independent cone degeneration (Zencak et al., 2013). This is consistent with our overall findings suggesting cell-autonomous degeneration of cones in *Bmi1*$^{-/-}$ mice.

Our genetic studies further revealed that *Chk2* or *p53* deletion could improve initial cone number in *Bmi1*$^{-/-}$ mice but not the outer segment morphology and progressive degeneration. This is consistent with previous work showing that in *Bmi1*$^{-/-}$ mice, *Chk2* inactivation improved several pathologies and lifespan, and that *p53* inactivation prevented cortical neuron cell death (Chatoo et al., 2009; Liu et al., 2009). However, the function of Bmi1 and p53 in the mouse retina, and especially in cones, are apparently not identical as in cortical neurons. For example, although p53-dependent apoptotic cell death was predominant in *Bmi1*$^{-/-}$ cortical neurons, it was not in *Bmi1*$^{-/-}$ cones, where we observed Rip3-associated necroptosis. At the molecular level, expression of the phase II antioxidant genes *Nqo1* and *Gsta1* was also not reduced in *Bmi1*$^{-/-}$ retinas, in contrast to *Bmi1*$^{-/-}$ cortices. Because these are direct p53 targets, this suggests that p53 activity is distinct between the retina and brain in the context of *Bmi1* deficiency.

In recent years, necroptosis was revealed as an alternative mechanism of cell death in many pathological contexts (Linkermann and Green, 2014; Vandenabeele et al., 2010). In the retina, it was found that cones, not rods, are especially vulnerable to necroptosis. For example, mutation in the cone-specific gene *PDE6C* results in achromaptosis, a disease characterized by cone degeneration (Chang et al., 2009). Using a zebrafish mutant for *pde6c*, it was shown that cone degeneration is mediated by rip1 and rip3, and that inhibition of necroptosis with necrostatin-1 can delay the disease process. Interestingly, necroptosis also operates in cones even when cone cell death is non cell-autonomous, such as in *rd10* mice carrying a mutation in the rod-specific gene *pde6b* (Murakami et al., 2012). Here again, treatment of *rd10* mice with necrostatin-1 greatly improved the cone degeneration phenotype, bringing hope for a possible pharmaceutical treatment of cone degenerative diseases.

Using directed differentiation of hES cells as model of retinal development, we demonstrated that BMI1 is required for the terminal differentiation of human cones. By contrast, expression of retinal homeobox genes and of *Opn1sw* was not altered in *Bmi1*$^{-/-}$ mouse retinas, revealing possible inter-species differences. These differences might, however, be explained by compensation mechanisms operating during development in *Bmi1*$^{-/-}$ mice but not in *BMI1* knockdown cells or by the highly distinct experimental systems. For example, acute *Bmi1* inactivation in embryonic mouse cortical progenitors was shown to induce massive apoptosis, in sharp contrast to the situation observed in *Bmi1*$^{-/-}$ mice (Fasano et al., 2007). Taken in a broader context, it is interesting to note that mice conditionally deficient for *Ezh2*, the catalytic subunit of the PRC2, which tri-methylates histone H3 at lysine 27, show reduced RPC proliferation and increased apoptosis, but not post natal degeneration of retinal neurons (Zhang et al., 2015). Likewise, mice deficient for *G9a* (*Ehmt2*), which di-methylates histone H3 at lysine 9, show increased RPC apoptosis and persistent cell proliferation (Katoh et al., 2012). In both cases, inefficient repression of non-retinal genes such as *Six1* (*Ezh2* knockout) or RPC genes such as *Hes1*, *Chx10* and *Lhx2* (*G9a* knockout) perturbs the terminal differentiation of retinal cells. These two situations are thus similar

Development (2016) 143, 1571-1584 doi:10.1242/dev.125351

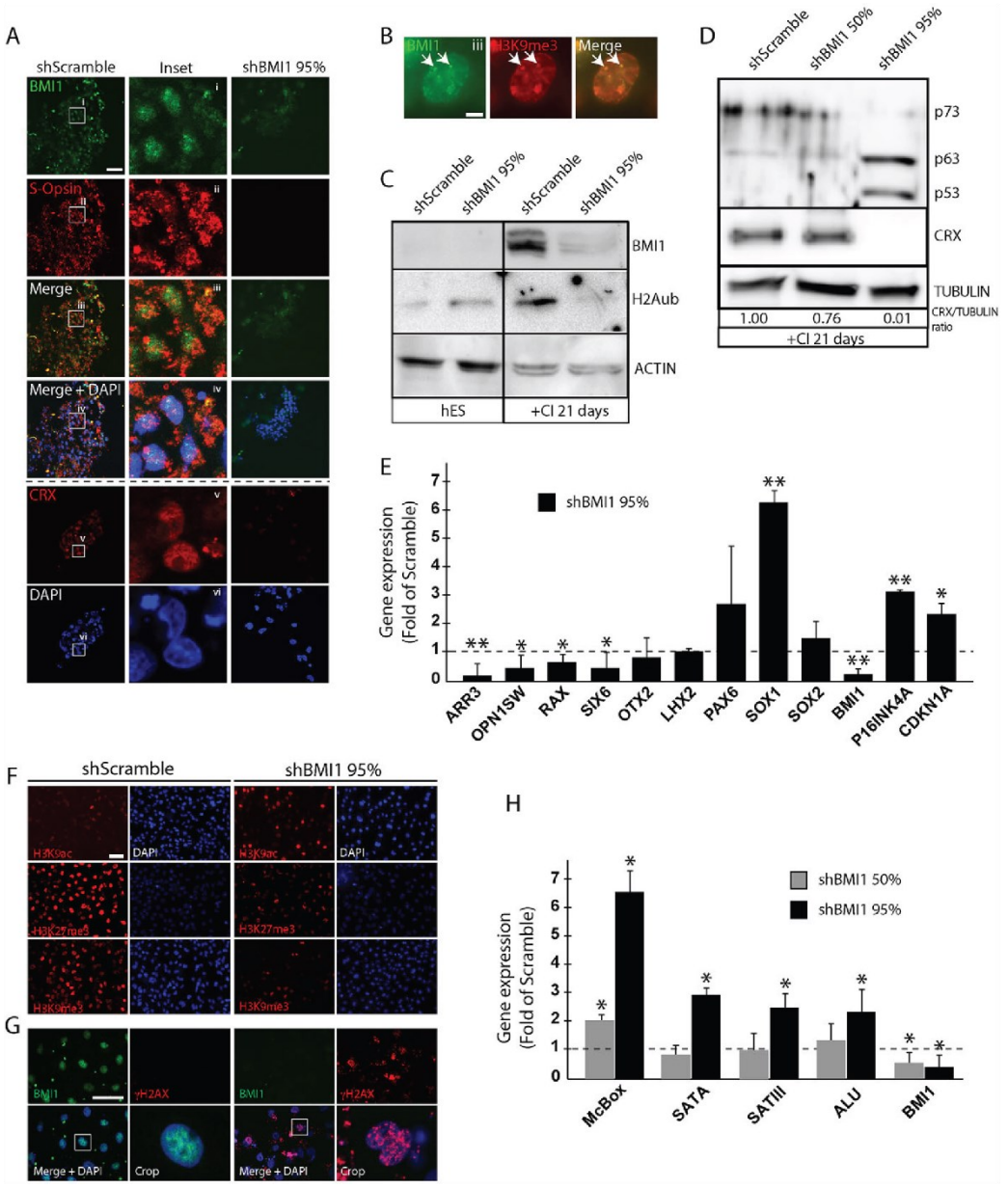


Fig. 9. BMI1 is required for terminal differentiation, heterochromatin compaction and genomic stability of human cones. (A,B) IF analysis of hES cells differentiated for 21 days toward the cone photoreceptor cell lineage. (A) In controls (shSrcamble), BMI1 is expressed in the nucleus of S-opsin^{+} cells. In shBMI1 cells, BMI1, S-opsin and CRX expression were highly reduced when compared with controls. (B) Colocalization of BMI1 and H3K9me3 in cones (arrows). (C,D) Western blot analyses of hES cells and *in vitro* differentiated cones (+CI 21 days) infected with shScramble or shBMI1 viruses. Note that BMI1 knockdown affects histone H2Aub levels only in cones; BMI1 knockdown in human cones was accompanied by induction of *p53*. Quantification of CRX levels normalized to tubulin; note the dose-dependent reduction in CRX expression in BMI1 knockdown cells. (E,H) Quantitative RT-PCR analyses of *in vitro* differentiated cones for retinal markers (E) and repeat DNA sequences (H). BMI1 was used as negative control and values are expressed as fold of Scramble value. (F,G) IF analyses of hES cell-derived cones (day 21). Scale bars: 20 µm in A,F,G; 2 µm in B. All values are means±s.e.m. *$P\leq0.05$; **$P\leq0.01$; Student's *t*-test.

Development (2016) 143, 1571-1584 doi:10.1242/dev.125351

to what we have observed in BMI1 deficient cones, where expression of *SOX1* was abnormally upregulated. At the level of the chromatin, it was demonstrated that rod photoreceptors of mice deficient for all three linker histone H1 genes have reduced chromatin condensation and increased nuclear diameter (Popova et al., 2013), a phenotype similar to that found in *Bmi1*$^{-/-}$ mouse photoreceptors and BMI1-deficient human cones. The results suggesting that BMI1 functions to prevent heterochromatin loss and expression of tandem repeats in photoreceptors are also novel findings. Since constitutive heterochromatin is the most instable portion of the mammalian genome (Bhaskara et al., 2010; Kappes et al., 2011; Larson et al., 2012; Peng and Karpen, 2009; Rowe et al., 2010), this could explain part of the genomic instability phenotype observed in BMI1 deficient cells (Chagraoui et al., 2011; Facchino et al., 2010; Ismail et al., 2010).

In conclusion, we demonstrated that Bmi1 expression in the retina is not required for retinal cell type genesis but is important to prevent bipolar neuron and cone photoreceptor degeneration during post natal development. Retinal cell death in *Bmi1*$^{-/-}$ mice was mediated by the activity of Chk2 and p53 on the one hand, and by Rip3-associated necroptosis on the other. Whether necroptosis is directly repressed by Bmi1 or indirectly mediated by the activity of p53 or Chk2 remains to be elucidated. Using an *in vitro* model of human cone development, we further demonstrated that BMI1 is required for the terminal differentiation of cones and maintenance of their genomic integrity. Further experiments, through inactivation of BMI1 in terminally differentiated cones, should address whether BMI1 is important to prevent human cone photoreceptor degeneration and/or maintenance of cell type identity.

MATERIALS AND METHODS

Mice and human tissues

Mice knockout for *p53*, *p16*^{INK4a} or *p16*^{INK4a}/*p19*^{ARF} were obtained from the Jackson Laboratory. Wild-type mice from the C57BL/6 genetic background were obtained from Charles River, Saint-Constant, Canada. Fresh human eyes were provided by the eye bank of Maisonneuve-Rosemont Hospital.

The Animal Care Committee of the Maisonneuve-Rosemont Hospital Research Centre approved the use of the animals in this study. Post mortem human eyes were provided by the Banque d'yeux du Québec du Centre Michel-Mathieu (http://www.maisonneuve-rosemont.org/pages/h/hopital/HMRCentreMichelMathieu.aspx) and were used with approbation of the Comité d'Éthique à la Recherche de l'Hôpital Maisonneuve-Rosemont. Human embryonic stem cells were used in accordance to Canadian Institute Health Research (CIHR) guidelines and approved by the Comité de Surveillance de la Recherche sur les Cellules Souches (CSRCS) of the CIHR.

Cell culture

The hES cell line H9 (WiCell) was cultured on BD matrigel-coated plate (BD bioscience) with a daily change of mTeSR medium according to the manufacturer's instructions (STEMCELL Technologies) (Thomson et al., 1998). The H9 hES cell line was first established on mouse embryonic fibroblasts (MEFs) and then cultured on Matrigel in mTeSR medium. For derivation of S-cones, undifferentiated hES cell colonies expanded at near confluence were cultured in DMEM-F12 medium supplemented with 2% B27, 1% N2, 1% NEAM (Life Technologies), 10 ng/ml IGF-1, 10 ng/ml FGF2 (Peprotech), 10 mg/ml heparin (Sigma) and 30 ng/ml Coco (R&D Systems) for 21 days. Retinal cells were cultured from eyes of WT and knockout mice as described in detail in supplementary Materials and Methods.

ERG recording and analysis

Electroretinograms (ERGs) were recorded on mice (WT, n=4; *Bmi1*$^{+/-}$, n=3; *Bmi1*$^{-/-}$, n=3) in photopic and scotopic conditions to assess the cone and rod systems activity, respectively, according to a previously described procedure (Lavoie et al., 2014). Detailed information is available in supplementary Materials and Methods.

Immunohistochemistry and immunofluorescence

Tissues were fixed by immersion in 4% paraformaldehyde (PFA)/3% sucrose in PBS, pH 7.4 for 1 h at room temperature. Samples were washed three times in PBS, cryoprotected in PBS/30% sucrose and frozen in Ctyomatrix embedding medium (CEM) (Thermo Shandon) or in Tissue-Tek optimum cutting temperature (O.C.T.) compound (Sakura Finetek). Otherwise, tissues were fixed in 10% buffered formalin and embedded in paraffin according to standard protocols. 5- to 12-µm-thick sections were mounted on Super-Frost glass slides (Fisher Scientific) and processed for immunofluorescence or immunohistochemistry staining. Detailed information on the experimental procedure and antibodies used is available in supplementary Materials and Methods.

Retinal dissection and eye orientation

Eyes were labeled in the dorsal pole of the cornea by puncturing with a needle tip. Eyes were extracted, fixed in 4% PFA overnight and stored in PBS. The retinas were marked by a small incision following the respective mark on the cornea. Retinas were then dissected as flattened whole mounts by four radial cuts and processed for IF using standard procedures, as describe above. Retinas were mounted between two coverslips that were attached to a slide by scotch tape, allowing flip-flopping of the retinas for analysis by microscope. For preparation of the blocks for cryosectioning, eyes were marked as described above, and oriented in blocks to maintain the eye polarity.

Quantification of retinal cell types

To collect sections from the ventral retina, blocks were trimmed up to the optic nerve and sections collected. For quantification, three successive images were taken on the nasal side from a set distance from the optic nerve. At least three different mice were used for each genotype (1 eye per animal), except for the *p53*$^{-/-}$/*Bmi1*$^{-/-}$ mouse where only two eyes were available. Positive cells were counted manually using ImageJ software (NIH). For IF analysis, single analysis of green channel (488) and red channel (568) were done to prevent filter overlap. To assess the general distribution of cones, whole mounts were photographed with a 10× objective using a Zeiss microscope (Observer.Z1) equipped with computer-driven motorized stage (VEXTA stepping motor), and individual frames were tiled to reconstruct the whole mounts (about 36 images/retina) using AxioVision 4.8 software. For quantification of IF images, photos were taken with a confocal microscope equipped with a 60× objective. Three images were taken per sample, where n=4 eyes per condition, at the same distance from the optic nerve in the naso-ventral portion of the retina. Photoreceptors labeled with PNA or with S-opsin and PNA were counted manually using ImageJ software. Quantifications of IHC experiments were performed in the same way but using images taken with a Zeiss microscope (Observer.Z1) equipped with a 25× objective.

Quantitative RT-PCR

All primers were designed to flank individual exons and tested by PCR in RT+ and RT− control extracts. Total RNA was isolated using TRIzol reagent (Invitrogen). Reverse transcription (RT) was performed using 1 µg total RNA and the MML-V reverse transcriptase (Invitrogen). Quantitative real-time PCR (qPCR) was performed using the Platinum SYBR Green SuperMix (Invitrogen) and a real-time PCR apparatus (ABI Prism 7000). GAPDH was used as an internal standard for data calibration. The $2^{-\Delta\Delta Ct}$ formula was used for the calculation of differential gene expression. All experiments were performed at least in triplicate. Primer sequences are available in supplementary Materials and Methods. Primer sets for repetitive sequences were as described previously (Zhu et al., 2011).

Lentiviral infection

The shRNA-expressing lentiviral plasmids were cotransfected with plasmids pCMVdR8.9 and pHCMVG into 293FT packaging cells using

Development (2016) 143, 1571-1584 doi:10.1242/dev.125351

Lipofectamine (Invitrogen) according to the manufacturer's instructions. Virus-containing media were collected, filtered and concentrated by ultracentrifugation. Viral titers were measured by serial dilution on 293T cells followed by microscopic analysis 48 h later. For viral transduction, lentiviral vectors were added to dissociated cells before plating. Hygromycin selection (150 µg/ml) was started 48 h later.

Western blot

Western blots of WT and *Bmi1* mutant mouse retinas were carried out using standard procedures with antibodies listed in supplementary Materials and Methods.

Statistical analyses

Statistical differences were analyzed using the Student's *t*-test for unpaired samples with equal s.d. using two-tailed *P*-value. For ERG experiments, analysis of variance were made by one-way ANOVA followed by Bonferroni's multiple comparisons test. For photoreceptor quantification in flat mount retinas, the analysis of variance was performed by two-way ANOVA followed by Tukey's multiple comparison test with 95% confidence. Data are presented as means±s.d. Values are representative of at least three experiments. The criterion for significance (*P*-value) was set as reported in the figures.

Acknowledgements

We thank M. Van Lohuizen for the *Bmi1*$^{+/-}$ mice and A. Swaroop for the *Nrl*$^{-/-}$ mouse eyes.

Competing interests

The authors declare no competing or financial interests.

Author contributions

Performed the experiments: A.B., V.P., M.A., W.C., A.F., R.H., S.Z. and J.L. Designed the experiments and wrote the manuscript: G.B., A.B., A.F., J.L. and M.H. Provided materials: N.M.

Funding

This work was supported by grants from the Natural Science and Engineering Research Council of Canada (NSERC); Antoine-Turmel Foundation for Macular Degeneration Research; and the Foundation Fighting Blindness Canada. V.P. was supported by a fellowship from NSERC. A.F. was supported by fellowships from Montreal University Molecular Biology Program and Réseau Vision du Québec.

Supplementary information

Supplementary information available online at
http://dev.biologists.org/lookup/suppl/doi:10.1242/dev.125351/-/DC1

References

Abdouh, M., Facchino, S., Chatoo, W., Balasingam, V., Ferreira, J. and Bernier, G. (2009). BMI1 sustains human glioblastoma multiforme stem cell renewal. *J. Neurosci.* **29**, 8884-8896.

Abdouh, M., Chatoo, W., El Hajjar, J., David, J., Ferreira, J. and Bernier, G. (2012). Bmi1 is down-regulated in the aging brain and displays antioxidant and protective activities in neurons. *PLoS ONE* **7**, e31870.

Avilion, A. A., Nicolis, S. K., Pevny, L. H., Perez, L., Vivian, N. and Lovell-Badge, R. (2003). Multipotent cell lineages in early mouse development depend on SOX2 function. *Genes Dev.* **17**, 126-140.

Belecky-Adams, T., Cook, B. and Adler, R. (1996). Correlations between terminal mitosis and differentiated fate of retinal precursor cells in vivo and in vitro: analysis with the "window-labeling" technique. *Dev. Biol.* **178**, 304-315.

Bernier, G., Panitz, F., Zhou, X., Hollemann, T., Gruss, P. and Pieler, T. (2000). Expanded retina territory by midbrain transformation upon overexpression of Six6 (Optx2) in Xenopus embryos. *Mech. Dev.* **93**, 59-69.

Bhaskara, S., Knutson, S. K., Jiang, G., Chandrasekharan, M. B., Wilson, A. J., Zheng, S., Yenamandra, A., Locke, K., Yuan, J.-L., Bonine-Summers, A. R. et al. (2010). Hdac3 is essential for the maintenance of chromatin structure and genome stability. *Cancer Cell* **18**, 436-447.

Boyer, L. A., Plath, K., Zeitlinger, J., Brambrink, T., Medeiros, L. A., Lee, T. I., Levine, S. S., Wernig, M., Tajonar, A., Ray, M. K. et al. (2006). Polycomb complexes repress developmental regulators in murine embryonic stem cells. *Nature* **441**, 349-353.

Buchwald, G., van der Stoop, P., Weichenrieder, O., Perrakis, A., van Lohuizen, M. and Sixma, T. K. (2006). Structure and E3-ligase activity of the Ring–Ring complex of polycomb proteins Bmi1 and Ring1b. *EMBO J.* **25**, 2465-2474.

Cao, R., Tsukada, Y.-i. and Zhang, Y. (2005). Role of Bmi-1 and Ring1A in H2A ubiquitylation and Hox gene silencing. *Mol. Cell* **20**, 845-854.

Carl, M., Loosli, F. and Wittbrodt, J. (2002). Six3 inactivation reveals its essential role for the formation and patterning of the vertebrate eye. *Development* **129**, 4057-4063.

Cepko, C. L., Austin, C. P., Yang, X., Alexiades, M. and Ezzeddine, D. (1996). Cell fate determination in the vertebrate retina. *Proc. Natl. Acad. Sci. USA* **93**, 589-595.

Chagraoui, J., Hebert, J., Girard, S. and Sauvageau, G. (2011). An anticlastogenic function for the Polycomb Group gene Bmi1. *Proc. Natl. Acad. Sci. USA* **108**, 5284-5289.

Chang, B., Grau, T., Dangel, S., Hurd, R., Jurklies, B., Sener, E. C., Andreasson, S., Dollfus, H., Baumann, B., Bolz, S. et al. (2009). A homologous genetic basis of the murine Cpfl1 mutant and human achromatopsia linked to mutations in the PDE6C gene. *Proc. Natl. Acad. Sci. USA* **106**, 19581-19586.

Chatoo, W., Abdouh, M., David, J., Champagne, M.-P., Ferreira, J., Rodier, F. and Bernier, G. (2009). The polycomb group gene Bmi1 regulates antioxidant defenses in neurons by repressing p53 pro-oxidant activity. *J. Neurosci.* **29**, 529-542.

Chatoo, W., Abdouh, M., Duparc, R.-H. and Bernier, G. (2010). Bmi1 distinguishes immature retinal progenitor/stem cells from the main progenitor cell population and is required for normal retinal development. *Stem Cells* **28**, 1412-1423.

Chatoo, W., Abdouh, M. and Bernier, G. (2011). P53 pro-oxidant activity in the central nervous system: implication in aging and neurodegenerative diseases. *Antioxid. Redox Signal.* **15**, 1729-1737.

Chen, S., Wang, Q.-L., Nie, Z., Sun, H., Lennon, G., Copeland, N. G., Gilbert, D. J., Jenkins, N. A. and Zack, D. J. (1997). Crx, a novel Otx-like paired-homeodomain protein, binds to and transactivates photoreceptor cell-specific genes. *Neuron* **19**, 1017-1030.

Chow, R. L., Altmann, C. R., Lang, R. A. and Hemmati-Brivanlou, A. (1999). Pax6 induces ectopic eyes in a vertebrate. *Development* **126**, 4213-4222.

Ellis, P., Fagan, B. M., Magness, S. T., Hutton, S., Taranova, O., Hayashi, S., McMahon, A., Rao, M. and Pevny, L. (2004). SOX2, a persistent marker for multipotential neural stem cells derived from embryonic stem cells, the embryo or the adult. *Dev. Neurosci.* **26**, 148-165.

Facchino, S., Abdouh, M., Chatoo, W. and Bernier, G. (2010). BMI1 confers radioresistance to normal and cancerous neural stem cells through recruitment of the DNA damage response machinery. *J. Neurosci.* **30**, 10096-10111.

Fasano, C. A., Dimos, J. T., Ivanova, N. B., Lowry, N., Lemischka, I. R. and Temple, S. (2007). shRNA knockdown of Bmi-1 reveals a critical role for p21-Rb pathway in NSC self-renewal during development. *Cell Stem Cell* **1**, 87-99.

Fodor, B. D., Shukeir, N., Reuter, G. and Jenuwein, T. (2010). Mammalian Su(var) genes in chromatin control. *Annu. Rev. Cell Dev. Biol.* **26**, 471-501.

Fortin, A., Cregan, S. P., MacLaurin, J. G., Kushwaha, N., Hickman, E. S., Thompson, C. S., Hakim, A., Albert, P. R., Cecconi, F., Helin, K. et al. (2001). APAF1 is a key transcriptional target for p53 in the regulation of neuronal cell death. *J. Cell Biol.* **155**, 207-216.

Freund, C. L., Gregory-Evans, C. Y., Furukawa, T., Papaioannou, M., Looser, J., Ploder, L., Bellingham, J., Ng, D., Herbrick, J.-A. S., Duncan, A. et al. (1997). Cone-rod dystrophy due to mutations in a novel photoreceptor-specific homeobox gene (CRX) essential for maintenance of the photoreceptor. *Cell* **91**, 543-553.

Freund, C. L., Wang, Q.-L., Chen, S., Muskat, B. L., Wiles, C. D., Sheffield, V. C., Jacobson, S. G., McInnes, R. R., Zack, D. J. and Stone, E. M. (1998). De novo mutations in the CRX homeobox gene associated with Leber congenital amaurosis. *Nat. Genet.* **18**, 311-312.

Furukawa, T., Morrow, E. M. and Cepko, C. L. (1997). Crx, a novel otx-like homeobox gene, shows photoreceptor-specific expression and regulates photoreceptor differentiation. *Cell* **91**, 531-541.

Furukawa, T., Morrow, E. M., Li, T., Davis, F. C. and Cepko, C. L. (1999). Retinopathy and attenuated circadian entrainment in Crx-deficient mice. *Nat. Genet.* **23**, 466-470.

Graham, V., Khudyakov, J., Ellis, P. and Pevny, L. (2003). Sox2 functions to maintain neural progenitor identity. *Neuron* **39**, 749-765.

Haverkamp, S. and Wässle, H. (2000). Immunocytochemical analysis of the mouse retina. *J. Comp. Neurol.* **424**, 1-23.

Hennig, A. K., Peng, G.-H. and Chen, S. (2013). Transcription coactivators p300 and CBP are necessary for photoreceptor-specific chromatin organization and gene expression. *PLoS ONE* **8**, e69721.

Ismail, I. H., Andrin, C., McDonald, D. and Hendzel, M. J. (2010). BMI1-mediated histone ubiquitylation promotes DNA double-strand break repair. *J. Cell Biol.* **191**, 45-60.

Jacobs, J. J., Kieboom, K., Marino, S., DePinho, R. A. and van Lohuizen, M. (1999). The oncogene and Polycomb-group gene bmi-1 regulates cell proliferation and senescence through the ink4a locus. *Nature* **397**, 164-168.

Jacobs, W. B., Kaplan, D. R. and Miller, F. D. (2006). The p53 family in nervous system development and disease. *J. Neurochem.* **97**, 1571-1584.

Jadhav, A. P., Mason, H. A. and Cepko, C. L. (2006). Notch 1 inhibits photoreceptor production in the developing mammalian retina. *Development* **133**, 913-923.

Jia, L., Oh, E. C., Ng, L., Srinivas, M., Brooks, M., Swaroop, A. and Forrest, D. (2009). Retinoid-related orphan nuclear receptor RORbeta is an early-acting factor in rod photoreceptor development. *Proc. Natl. Acad. Sci. USA* **106**, 17534-17539.

Kappes, F., Waldmann, T., Mathew, V., Yu, J., Zhang, L., Khodadoust, M. S., Chinnaiyan, A. M., Luger, K., Erhardt, S., Schneider, R. et al. (2011). The DEK oncoprotein is a Su(var) that is essential to heterochromatin integrity. *Genes Dev.* **25**, 673-678.

Karimi, M. M., Goyal, P., Maksakova, I. A., Bilenky, M., Leung, D., Tang, J. X., Shinkai, Y., Mager, D. L., Jones, S., Hirst, M. et al. (2011). DNA methylation and SETDB1/H3K9me3 regulate predominantly distinct sets of genes, retroelements, and chimeric transcripts in mESCs. *Cell Stem Cell* **8**, 676-687.

Katoh, K., Yamazaki, R., Onishi, A., Sanuki, R. and Furukawa, T. (2012). G9a histone methyltransferase activity in retinal progenitors is essential for proper differentiation and survival of mouse retinal cells. *J. Neurosci.* **32**, 17658-17670.

Lagutin, O. V., Zhu, C. C., Kobayashi, D., Topczewski, J., Shimamura, K., Puelles, L., Russell, H. R. C., McKinnon, P. J., Solnica-Krezel, L. and Oliver, G. (2003). Six3 repression of Wnt signaling in the anterior neuroectoderm is essential for vertebrate forebrain development. *Genes Dev.* **17**, 368-379.

Larson, K., Yan, S. J., Tsurumi, A., Liu, J., Zhou, J., Gaur, K., Guo, D., Eickbush, T. H. and Li, W. X. (2012). Heterochromatin formation promotes longevity and represses ribosomal RNA synthesis. *PLoS Genet.* **8**, e1002473.

Lavoie, J., Illiano, P., Sotnikova, T. D., Gainetdinov, R. R., Beaulieu, J. M. and Hebert, M. (2014). The electroretinogram as a biomarker of central dopamine and serotonin: potential relevance to psychiatric disorders. *Biol. Psychiatry* **75**, 479-486.

Li, Z., Cao, R., Wang, M., Myers, M. P., Zhang, Y. and Xu, R.-M. (2006). Structure of a Bmi-1-Ring1B polycomb group ubiquitin ligase complex. *J. Biol. Chem.* **281**, 20643-20649.

Linkermann, A. and Green, D. R. (2014). Necroptosis. *N. Engl. J. Med.* **370**, 455-465.

Liu, J., Cao, L., Chen, J., Song, S., Lee, I. H., Quijano, C., Liu, H., Keyvanfar, K., Chen, H., Cao, L.-Y. et al. (2009). Bmi1 regulates mitochondrial function and the DNA damage response pathway. *Nature* **459**, 387-392.

Loosli, F., Winkler, S. and Wittbrodt, J. (1999). Six3 overexpression initiates the formation of ectopic retina. *Genes Dev.* **13**, 649-654.

Marigo, V. (2007). Programmed cell death in retinal degeneration: targeting apoptosis in photoreceptors as potential therapy for retinal degeneration. *Cell Cycle* **6**, 652-655.

Marquardt, T. (2003). Transcriptional control of neuronal diversification in the retina. *Prog. Retin. Eye Res.* **22**, 567-577.

Marquardt, T., Ashery-Padan, R., Andrejewski, N., Scardigli, R., Guillemot, F. and Gruss, P. (2001). Pax6 is required for the multipotent state of retinal progenitor cells. *Cell* **105**, 43-55.

Mathers, P. H., Grinberg, A., Mahon, K. A. and Jamrich, M. (1997). The Rx homeobox gene is essential for vertebrate eye development. *Nature* **387**, 603-607.

Mears, A. J., Kondo, M., Swain, K. S., Takada, Y., Bush, R. A., Saunders, T. L., Sieving, P. A. and Swaroop, A. (2001). Nrl is required for rod photoreceptor development. *Nat. Genet.* **29**, 447-452.

Molofsky, A. V., He, S., Bydon, M., Morrison, S. J. and Pardal, R. (2005). Bmi-1 promotes neural stem cell self-renewal and neural development but not mouse growth and survival by repressing the p16Ink4a and p19Arf senescence pathways. *Genes Dev.* **19**, 1432-1437.

Murakami, Y., Matsumoto, H., Roh, M., Suzuki, J., Hisatomi, T., Ikeda, Y., Miller, J. W. and Vavvas, D. G. (2012). Receptor interacting protein kinase mediates necrotic cone but not rod cell death in a mouse model of inherited degeneration. *Proc. Natl. Acad. Sci. USA* **109**, 14598-14603.

Ng, L., Hurley, J. B., Dierks, B., Srinivas, M., Saltó, C., Vennström, B., Reh, T. A. and Forrest, D. (2001). A thyroid hormone receptor that is required for the development of green cone photoreceptors. *Nat. Genet.* **27**, 94-98.

Nishida, A., Furukawa, A., Koike, C., Tano, Y., Aizawa, S., Matsuo, I. and Furukawa, T. (2003). Otx2 homeobox gene controls retinal photoreceptor cell fate and pineal gland development. *Nat. Neurosci.* **6**, 1255-1263.

Ortín-Martínez, A., Nadal-Nicolás, F. M., Jiménez-López, M., Alburquerque-Béjar, J. J., Nieto-López, L., García-Ayuso, D., Villegas-Pérez, M. P., Vidal-Sanz, M. and Agudo-Barriuso, M. (2014). Number and distribution of mouse retinal cone photoreceptors: differences between an albino (Swiss) and a pigmented (C57/BL6) strain. *PLoS ONE* **9**, e102392.

Peng, J. C. and Karpen, G. H. (2009). Heterochromatic genome stability requires regulators of histone H3 K9 methylation. *PLoS Genet.* **5**, e1000435.

Popova, E. Y., Grigoryev, S. A., Fan, Y., Skoultchi, A. I., Zhang, S. S. and Barnstable, C. J. (2013). Developmentally regulated linker histone H1c promotes heterochromatin condensation and mediates structural integrity of rod photoreceptors in mouse retina. *J. Biol. Chem.* **288**, 17895-17907.

Porter, F. D., Drago, J., Xu, Y., Cheema, S. S., Wassif, C., Huang, S. P., Lee, E., Grinberg, A., Massalas, J. S., Bodine, D. et al. (1997). Lhx2, a LIM homeobox gene, is required for eye, forebrain, and definitive erythrocyte development. *Development* **124**, 2935-2944.

Reh, T. A. and Kljavin, I. J. (1989). Age of differentiation determines rat retinal germinal cell phenotype: induction of differentiation by dissociation. *J. Neurosci.* **9**, 4179-4189.

Rogakou, E. P., Pilch, D. R., Orr, A. H., Ivanova, V. S. and Bonner, W. M. (1998). DNA double-stranded breaks induce histone H2AX phosphorylation on serine 139. *J. Biol. Chem.* **273**, 5858-5868.

Rowe, H. M., Jakobsson, J., Mesnard, D., Rougemont, J., Reynard, S., Aktas, T., Maillard, P. V., Layard-Liesching, H., Verp, S., Marquis, J. et al. (2010). KAP1 controls endogenous retroviruses in embryonic stem cells. *Nature* **463**, 237-240.

Sancho-Pelluz, J., Arango-Gonzalez, B., Kustermann, S., Romero, F. J., van Veen, T., Zrenner, E., Ekström, P. and Paquet-Durand, F. (2008). Photoreceptor cell death mechanisms in inherited retinal degeneration. *Mol. Neurobiol.* **38**, 253-269.

Sauvageau, M. and Sauvageau, G. (2010). Polycomb group proteins: multi-faceted regulators of somatic stem cells and cancer. *Cell Stem Cell* **7**, 299-313.

Sharpless, N. E., Ramsey, M. R., Balasubramanian, P., Castrillon, D. H. and DePinho, R. A. (2004). The differential impact of p16(INK4a) or p19(ARF) deficiency on cell growth and tumorigenesis. *Oncogene* **23**, 379-385.

Sherr, C. J. (2001). The INK4a/ARF network in tumour suppression. *Nat. Rev. Mol. Cell Biol.* **2**, 731-737.

Solovei, I., Kreysing, M., Lanctôt, C., Kösem, S., Peichl, L., Cremer, T., Guck, J. and Joffe, B. (2009). Nuclear architecture of rod photoreceptor cells adapts to vision in mammalian evolution. *Cell* **137**, 356-368.

Swaroop, A., Wang, Q.-L., Wu, W., Cook, J., Coats, C., Xu, S., Chen, S., Zack, D. J. and Sieving, P. A. (1999). Leber congenital amaurosis caused by a homozygous mutation (R90W) in the homeodomain of the retinal transcription factor CRX: direct evidence for the involvement of CRX in the development of photoreceptor function. *Hum. Mol. Genet.* **8**, 299-305.

Swaroop, A., Kim, D. and Forrest, D. (2010). Transcriptional regulation of photoreceptor development and homeostasis in the mammalian retina. *Nat. Rev. Neurosci.* **11**, 563-576.

Taranova, O. V., Magness, S. T., Fagan, B. M., Wu, Y., Surzenko, N., Hutton, S. R. and Pevny, L. H. (2006). SOX2 is a dose-dependent regulator of retinal neural progenitor competence. *Genes Dev.* **20**, 1187-1202.

Thomson, J. A., Itskovitz-Eldor, J., Shapiro, S. S., Waknitz, M. A., Swiergiel, J. J., Marshall, V. S. and Jones, J. M. (1998). Embryonic stem cell lines derived from human blastocysts. *Science* **282**, 1145-1147.

Valk-Lingbeek, M. E., Bruggeman, S. W. M. and van Lohuizen, M. (2004). Stem cells and cancer; the polycomb connection. *Cell* **118**, 409-418.

van der Lugt, N. M., Domen, J., Linders, K., van Roon, M., Robanus-Maandag, E., te Riele, H., van der Valk, M., Deschamps, J., Sofroniew, M., van Lohuizen, M. et al. (1994). Posterior transformation, neurological abnormalities, and severe hematopoietic defects in mice with a targeted deletion of the bmi-1 proto-oncogene. *Genes Dev.* **8**, 757-769.

Vandenabeele, P., Galluzzi, L., Vanden Berghe, T. and Kroemer, G. (2010). Molecular mechanisms of necroptosis: an ordered cellular explosion. *Nat. Rev. Mol. Cell Biol.* **11**, 700-714.

Viringipurampeer, I. A., Shan, X., Gregory-Evans, K., Zhang, J. P., Mohammadi, Z. and Gregory-Evans, C. Y. (2014). Rip3 knockdown rescues photoreceptor cell death in blind pde6c zebrafish. *Cell Death Differ.* **21**, 665-675.

Wang, H., Wang, L., Erdjument-Bromage, H., Vidal, M., Tempst, P., Jones, R. S. and Zhang, Y. (2004). Role of histone H2A ubiquitination in Polycomb silencing. *Nature* **431**, 873-878.

Watanabe, T. and Raff, M. C. (1990). Rod photoreceptor development in vitro: intrinsic properties of proliferating neuroepithelial cells change as development proceeds in the rat retina. *Neuron* **4**, 461-467.

Yan, Y., Yang, D., Zarnowska, E. D., Du, Z., Werbel, B., Valliere, C., Pearce, R. A., Thomson, J. A. and Zhang, S.-C. (2005). Directed differentiation of dopaminergic neuronal subtypes from human embryonic stem cells. *Stem Cells* **23**, 781-790.

Yanagi, Y., Takezawa, S. and Kato, S. (2002). Distinct functions of photoreceptor cell-specific nuclear receptor, thyroid hormone receptor beta2 and CRX in one photoreceptor development. *Invest. Ophthalmol. Vis. Sci.* **43**, 3489-3494.

Yaron, O., Farhy, C., Marquardt, T., Applebury, M. and Ashery-Padan, R. (2006). Notch1 functions to suppress cone-photoreceptor fate specification in the developing mouse retina. *Development* **133**, 1367-1378.

Zencak, D., Schouwey, K., Chen, D., Ekstrom, P., Tanger, E., Bremner, R., van Lohuizen, M. and Arsenijevic, Y. (2013). Retinal degeneration depends on Bmi1 function and reactivation of cell cycle proteins. *Proc. Natl. Acad. Sci. USA* **110**, E593-E601.

Zhang, J., Taylor, R. J., La Torre, A., Wilken, M. S., Cox, K. E., Reh, T. A. and Vetter, M. L. (2015). Ezh2 maintains retinal progenitor proliferation, transcriptional integrity, and the timing of late differentiation. *Dev. Biol.* **403**, 128-138.

Zhou, S., Flamier, A., Abdouh, M., Tetreault, N., Barabino, A., Wadhwa, S. and Bernier, G. (2015). Differentiation of human embryonic stem cells into cone photoreceptors through simultaneous inhibition of BMP, TGFbeta and Wnt signaling. *Development* **142**, 3294-3306.

Zhu, Q., Pao, G. M., Huynh, A. M., Suh, H., Tonnu, N., Nederlof, P. M., Gage, F. H. and Verma, I. M. (2011). BRCA1 tumour suppression occurs via heterochromatin-mediated silencing. *Nature* **477**, 179-184.

SUPPLEMENTARY MATERIALS AND METHODS

ERG recording and analysis

A typical ERG trace is composed of the a-wave, which is a negative component originating from the photoreceptors, and the b-wave, which is a positive component generated by the bipolar and Müller cells complex stimulation. Two major parameters can be derived from these waves, namely the amplitude (μV) and the implicit time (ms). By convention, the amplitude of the a-wave is measured from the baseline to trough, and the b-wave amplitude is measured from the trough of the a-wave to the peak of the b-wave. The implicit time of the waves represents the number of milliseconds at which the maximal amplitude is reached. For each of the four parameters, a photopic and scotopic luminance response function (LRF) was generated, where the values of the parameters were plotted against flash luminance. Repeated measures analyses of variance with Bonferroni correction were performed to assess the difference between the three genotypes for each one of the parameters. All analyses were conducted using SPSS for Windows, version 22.0.

Immunohistochemistry and immunofluorescence

For immunofluorescence labeling, sections were incubated overnight with primary antibody solutions at 4ºC in a humidified chamber. After three washes in PBS, sections were incubated with secondary antibodies for 1 h at room temperature. Slides were mounted on coverslips in DAPI-containing mounting medium (Vector Laboratories, CA). For immunohistochemistry labeling, slices were analyzed by using the Vectastain® ABC kit (Vector) according to the manufacturer instructions. Peroxidase substrates used are the Vector® VIP (Pink) (Vector), and DAB (brown) (Sigma). Observations were made under a fluorescence microscope (Leica DMRE, Leica Microsystems) and images were captured with a digital camera (Retiga EX; QIMAGING; with OpenLab, ver.3.1.1 software; Open-Lab, Canada). Confocal microscopy analyses were performed using 60x objectives with an IX81 confocal microscope (Olympus, Richmond Hill, Canada), and images were obtained with Fluoview software version 3.1 (Olympus). 3D reconstructions were obtained with Fluoview software version 3.1 from 18-25 z-stack image. Primary antibodies used in this study are: sheep anti-Chx10 (1:250, Exalpha Biologicals), rabbit anti-Pax6 (1:500, Chemicon), mouse anti-Bmi1 (1:200, US Biological), rabbit anti-Bmi1 (1:150, US Biological), mouse anti-Syntaxin (1:200, Sigma), mouse anti-4D2 (Rhodopsin) (1:100, R. Molday, UBC), mouse anti-Gad65 (1:300, BD Pharmingen), rabbit anti-CORD2 (CRX) (1:300, Abcam), rabbit anti-S-Opsin (1:200, Invitrogen), rabbit anti-M-Opsin (1:100, Chemicon), Rip3 (1:250, Santa Cruz), rabbit anti-H3K9me3 (1:500, Abcam), rabbit anti-H3K9ac (1:300, Cell Signaling), rabbit anti-H3K27me3 (1:300, Cell Signaling), mouse anti-glutamine synthetase (GS) (1:100, Chemicon), rabbit anti-Pkcα

(1:500, Sigma), rabbit anti-Calbidin (1:500, Chemicon), rabbit anti-Recoverin (1:1000, Millipore), mouse anti-γH2Ax (1:250, Millipore), rabbit anti-Cralbp (1:500, kindly given by Dr Saari's lab). Secondary antibodies are: donkey AlexaFluor488-conjugated anti-mouse (1:1000, Life Technologies), donkey AlexaFluor488-conjugated anti-rabbit (1:1000, Life Technologies), goat AlexaFluor594-conjugated anti-mouse IgM (1:1000, Invitrogen), donkey AlexaFluor633-conjugated anti-sheep (1:1000, Molecular Probes), goat AlexaFluor647-conjugated anti-mouse (1:1000, Life Technologies) goat AlexaFluor texas red-conjugated anti-rabbit (1:1000, Life Technologies), donkey FITC-conjugated anti-mouse (1:300, Chemicon), goat FITC-conjugated anti-rat (1:300, Caltag Laboratories), donkey Rhodamine-conjugated anti-rabbit (1:300, Chemicon). Fluorescein labeled Peanut Agglutinin (PNA) (1:200, Vector laboratories) was used to stain the outer segment of PRs. Superoxide production was measured by red fluorescence intensity as described by standard protocol indicated by manufacturer (MitoSoxRed mitochondrial superoxide indicator, Molecular Probes, U.S.).

Quantitative RT-PCR

Mouse primer sets used are:

Tnf (F) 5' AAAATTCGAGTGACAAGCCTGTAG 3'; Tnf (R) 5' CCCTTGAAGAGAACCTGGGAGTAG 3'; Tnfrsf1a (F) 5' GCCGGATATGGGCATGAAGC 3'; Tnfrsf1a (R) 5' TGTCTCAGCCCTCACTTGAC 3'; Ripk1 (F) 5' TGTCATCTAGCGGGAGGTTG 3'; Ripk1 (R) 5' TCACCACTCGACTGTGTCTCAG 3'; Ripk3 (F) 5' CTCCGTGCCTTGACCTACTG 3'; Ripk3 (R) 5' AACCATAGCCTTCACCTCCC3'; Bmi1 (F) 5'-GGAGACCAGCAAGTATTGTCCTATTTG-3', Bmi1 (R) 5'-CTTACGATGCCCAGCAGCAATG-3'; $p16^{Ink4a}$ (F) 5'-CAACGCCCCGAACTCTTTC-3', $p16^{Ink4a}$ (R) 5'-GCAGAAGAGCTGCTACGTGAAC-3'; $p19^{Arf}$ (F) 5'-GGCTAGAGAGGATCTTGAGAAGAGG-3', $p19^{Arf}$ (R) 5'-GCCCATCATCATCACCTGGTCCAGG-3'; Sox2 (F) 5'-TAAGGGTTCTTGCTGGGTTTT-3', Sox2 (R) 5'-AGACCACGAAAACGGTCTTG-3'; Lhx2 (F) 5'-GATCTCGCCTGGAAACAGAG -3', Lhx2 (R) 5'-TCGCTCAGTCCACAAAACTG -3'; Otx2 (F) 5'-AGAGGAGGTGGCACTGAAAA-3', Otx2 (R) 5'-TGACCTCCATTCTGCTGTTG-3'; Lpo (F) 5'-AGGTCTGTTGGCCAAGAATG-3', Lpo (R) 5'-ATGTTGATGGAAGCCAGGTC-3'; Apaf1 (F) 5'-TGCTCAGCGGATAAGAAGGT-3', Apaf1 (R) 5'TCCCAGAGCTTGAGGAAGAA-3'; Fas (F) 5'-AAACAAACTGCACCCTGACC-3', Fas (R) 5'CAACCATAGGCGATTTCTGG-3'; Nqo1 (F) 5'-TTCTCTGGCCGATTCAGAGT-3', Nqo1 (R) 5'GAGTGTGGCCAATGCTGTAA-3'; Gsta1 (F) 5'-CGCCACCAAATATGACCTCT-3', Gsta1 (R) 5'CCATGGCTCTTCAACACCTT-3';

Cyp24a1 (F) 5'-GGCGGAAGATGTGAGGAATA-3', Cyp24a1 (R) 5'GTTGTGAATGGCACACTTGG-3'; Duox2 (F) 5'-ACAAGGGGTGTATGCCTTTG-3', Duox2 (R) 5'-CACAGGTTGTGGTAGCGAAA-3'. Crx (F) 5'-CCTTCTGACAGCTCGGTGTT-3', Crx (R) 5'-CCACTTTCTGAAGCCTGGAG)-3'; Sesn2 (F) 5'-CCTCCTTTGTGTTGTGCTGT-3', Sesn2 (R) 5'-ACGGTTCTCCATTTCCTCCT-3'; Opn1sw (F) 5'-CAGCCTTCATGGGATTTGTCT-3', Opn1sw (R) 5'-CAAAGAGGAAGTATCCGTGAC-3'; Rax (F) 5′-TGGGCTTTACCAAGGAAGACG-3', Rax (R) 5′-GGTAGCAGGGCCTAGTAGCTT-3'; Six6 (F) 5′-GCAAGTAGCCGGGGTATGTG-3', Six6 (R) 5′-CGACTCATTCTTGTTAAGGGCTT-3'; Nrl (F) 5′-CCCAGTCCCTTGGCTATGGA-3', Nrl (R) 5′-ACCGAGCTGTATGGTGTGGA-3'; Notch1 (F) 5′-GATGGCCTCAATGGGTACAAG-3', Notch1 (R) 5′-TCGTTGTTGTTGATGTCACAGT-3'; Pax6 (F) 5′-TGGCAAACAACCTGCCTATG-3', Pax6 (R) 5′-TGCACGAGTATGAGGAGGTCT-3'; Gapdh (F) 5′-AGGTCGGTGTGAACGGATTTG-3′, Gapdh (R) 5′-TGTAGACCATGTAGTTGAGGTCA-3';

Human primer sets used are:

PAX6 (F) 5'-AGATTTCAGAGCCCCATATTCG-3', PAX6 (R) 5'-CCATTTGGCCCTTCGATTAG-3'; ARR3 (F) 5'-CCCAGAGCTTTGCAGTAACC-3', ARR3 (R) 5'-CACAGGACACCATCAGGTTG-3'; SOX1 (F) 5'-AAAGTCAAAACGAGGCGAGA-3', SOX1 (R) 5'-AAGTGCTTGGACCTGCCTTA-3'; RAX (F) 5'-GGCAAGGTCAACCTACCAGA-3', RAX (R) 5'-GCTTCATGGAGGACACTTCC-3'; SIX6 (F) 5′- ACAGACTCCAGCAGCAGGTT-3′, SIX6 (R) 5′-AGATGTCGCACTCACTGTCG-3′; OPN1SW (F) 5′-TGTGCCTCTCTCCCTCATCT-3′, OPN1SW (R) 5′-GGCACGTAGCAGACACAGAA-3′; $p16^{INK4A}$ (F) 5′-GATCCAGGTGGGTAGAAGGTC-3′, $p16^{INK4A}$ (R) 5′-CCCCTGCAAACTTCGTCCT-3′; CDKN1A (F) 5′-CCGAAGTCAGTTCCTTGTGG-3′, CDKN1A (R) 5′-GTCGAAGTTCCATCGCTCAC3′; SOX2 (F) 5′-CACAACTCGGAGATCAGCAA-3′, SOX2 (R) 5′-CGGGGCCGGTATTTATAATC-3′; LHX2 (F) 5′-CCAAGGACTTGAAGCAGCTC-3′, LHX2 (R) 5′-TAAGAGGTTGCGCCTGAACT-3′; BMI1 (F) 5-AATCCCCACCTGATGTGTGT-3, BMI1 (R) 5-GCTGGTCTCCAGGTA ACGAA-3; GAPDH (F) 5-TCACCAGGGCTGCTTTTAAC-3′, GAPDH (R) 5′-ATCCACAGTCTTCTGGGTGG -3′.

Western blot

Total protein extracts were prepared in the Complete Mini protease inhibitor cocktail solution (Roche Diagnostics). Proteins contents were quantified using the Bradford reagent. Proteins were

resolved in Laemmli buffer by SDS-PAGE and transferred to a 0.2µm Nitrocellulose Blotting Membrane (BioRad) that was exposed to the primary antibodies: mouse anti-Bmi1 (1:800, Millipore), mouse anti-H2Aub (1:1000, Millipore), mouse anti- βActin (1:1000, abcam), p63, mouse anti-p73 (1:500, abcam), mouse anti-p53 (1:500, Santa Cruz Biotecnology), goat anti-CRX (1:500, Santa Cruz Biotecnology), mouse anti-α-Tubulin (1:1000, Sigma), anti-Rip3 (1:1000, Santa Cruz Biotecnology), anti-Bmi1 (1:500, abgent), S-Opsin (1:400, Santa Cruz Biotechnology) and histone H3 (1:1000, upstate). Membranes were treated with corresponding horseradish peroxidase-conjugated secondary antibodies (Sigma) and developed using the Immobilon Western (Millipore).

Cell cultures

Cultures of retinal cells were obtained by dissecting the mice eyes at P5 in 1X oxygenated HBSS (Life Technologies) in order to extract only the neural retina. Retinal cells were then re-suspended and incubated 10 minutes at 37 °C in the enzyme solution composed of 10ml 1X HBSS, 9.3 mg of Papain (Worthington), 1.6 mg of N-acetyl L-cysteine (Sigma), 0.5 mg of DNaseI (Roche) and 10µl EDTA 500mM (Fisher Scientific). After centrifugation cells were dissociated into Neurobasal^{tm} medium (Life technologies) with 0.02 µg/µl NGF (Invitrogen), 0.02 µg/µl BDNF (Invitrogen), 1% B27 (Invitrogen), 70 µg/ml gentamycin (Invitrogen), 1% fetal bovine serum (Wisent), 0.5% glucose (Sigma) and 10 µM Forskolin (Sigma). The cells were then spread and cultivated on coverslips treated with Poly-L-Lysine hydrobromide (Sigma) and BD Matrigel Matrix (BD Biosciences).

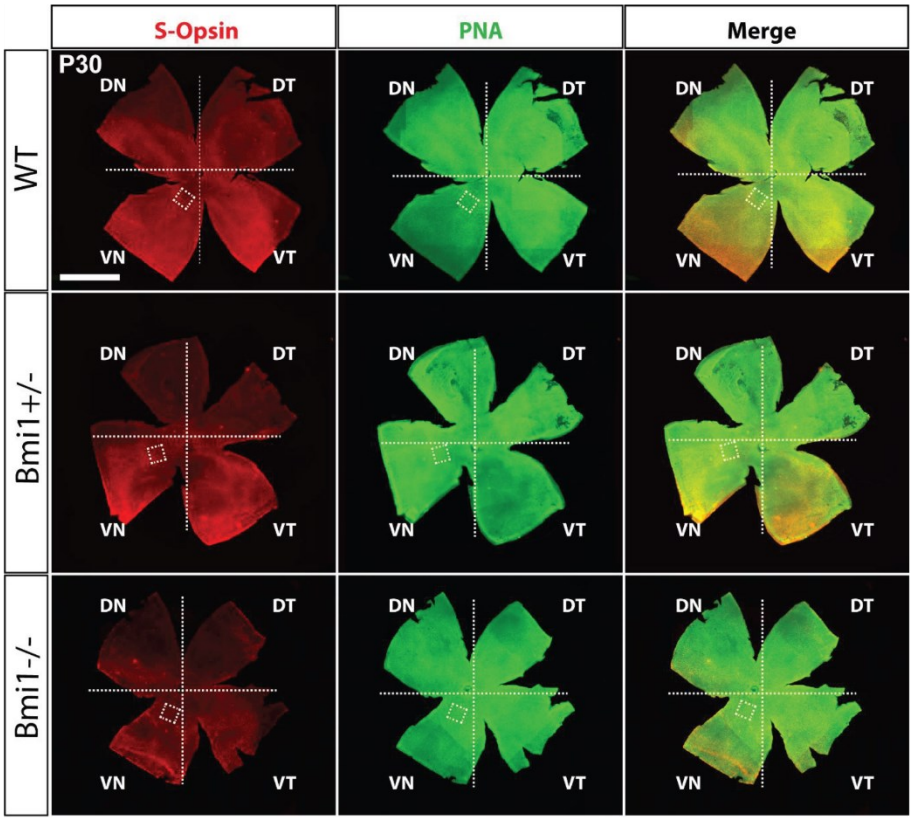


Figure S1. *Bmi1* is required for cone photoreceptor maintenance after birth

(A) Representative images from S-Opsin/PNA double-stained retinal flat mounts from WT, *Bmi1*$^{+/-}$ and *Bmi1*$^{-/-}$ mice at P30. Dorso-nasal (DN), dorso-temporal (DT), ventro-nasal (VN), ventro-temporal (VT). Scale bars: 1 mm.

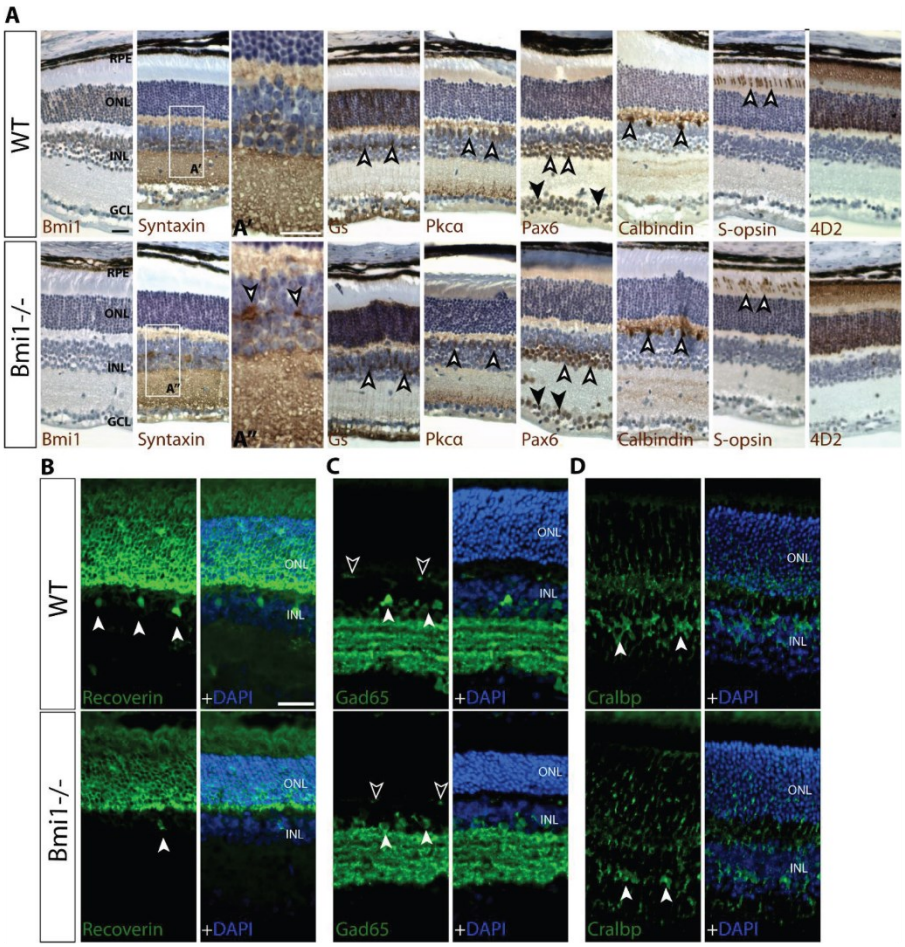


Figure S2. Histological anomalies in *Bmi1*$^{-/-}$ mouse retinas

(A, B, C, D) IHC and IF analyses of WT and *Bmi1*$^{-/-}$ retinas at P30 using specific markers for different retinal cell types. (A', A'') crop of the area indicated by the respective dashed rectangles. Amacrine cells' membrane (Syntaxin), müller glial cells (Gs and Cralbp), rods bipolar cells (Pkcα), amacrine (white-edged arrows) and ganglion cells (black arrows) (Pax6), horizontal cells (Calbindin), c-cones (S-Opsin), rod photoreceptors (4D2), T2 OFF and T8 ON cone bipolar cells (Recoverin), amacrine cells (white arrows) and horizontal cells (white-edged arrows) (Gad65). Retinal pigmented epithelium (RPE); outer nuclear layer (ONL); inner nuclear layer (INL); and ganglion cell layer (GCL). Scale bars: 40µm.

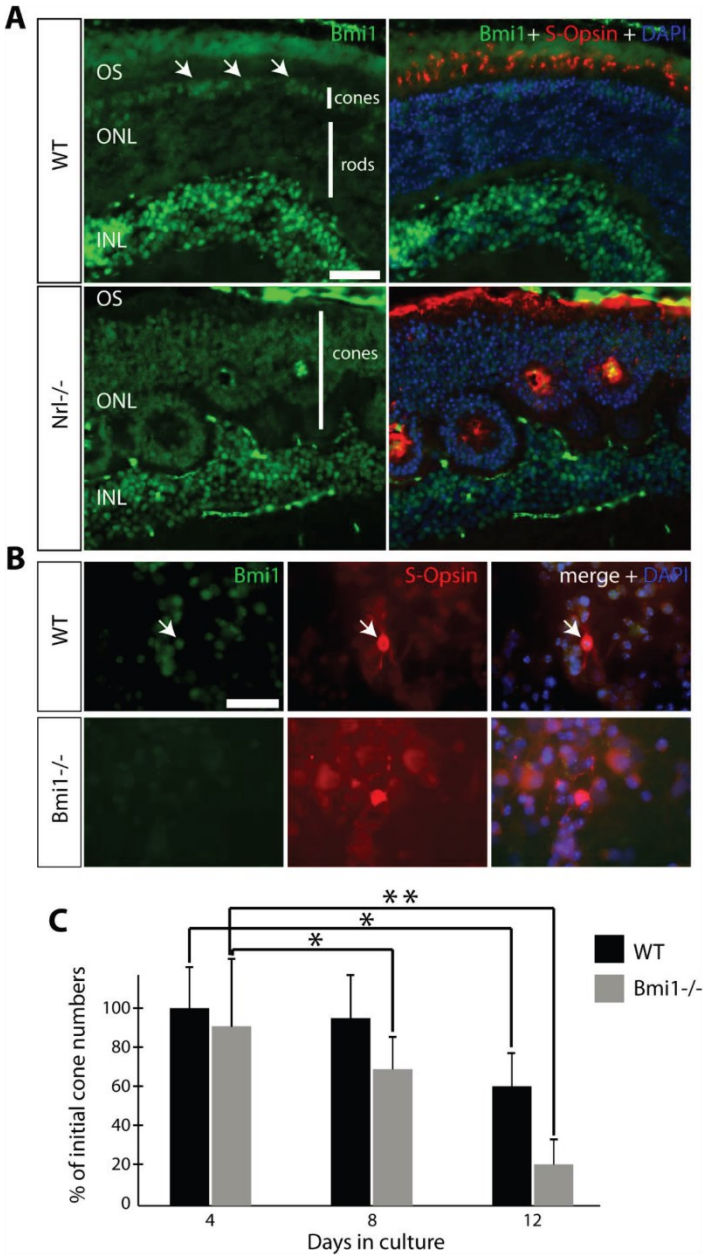


Figure S3. Bmi1 is preferentially expressed in mouse cone photoreceptors

(A) IF analysis of WT and *Nrl*$^{-/-}$ mouse retinal cryosections at P30. Arrows: Note the preferential expression of Bmi1 in neurons of the INL and in cone photoreceptors located in the ONL (white arrows). In contrast, Bmi1 is evenly distributed in the cone-only ONL of *Nrl*$^{-/-}$ mice. (B) IF analysis

of dissociated retinal cultures from WT and *Bmi1*$^{-/-}$ mice at P1 after 4 days *in vitro* (DIV). Note the expression of Bmi1 in WT cones labeled with S-Opsin, and the absence of Bmi1 in S-Opsin positive cells of *Bmi1*-mutants. (C) Quantification of S-Opsin positive cones in dissociated WT and *Bmi1*$^{-/-}$ retinal cultures, expressed as percentage of the initial cone number evaluated at 1 DIV. Outer nuclear layer (ONL); inner nuclear layer (INL); outer segment (OS). Scale bar: 50µm (A) and 20µm (B). WT n=4; *Bmi1*$^{-/-}$ n= 5. All values are mean ±SEM. (*) P ≤ 0.05; (**) ≤ 0.01; Student t-test.

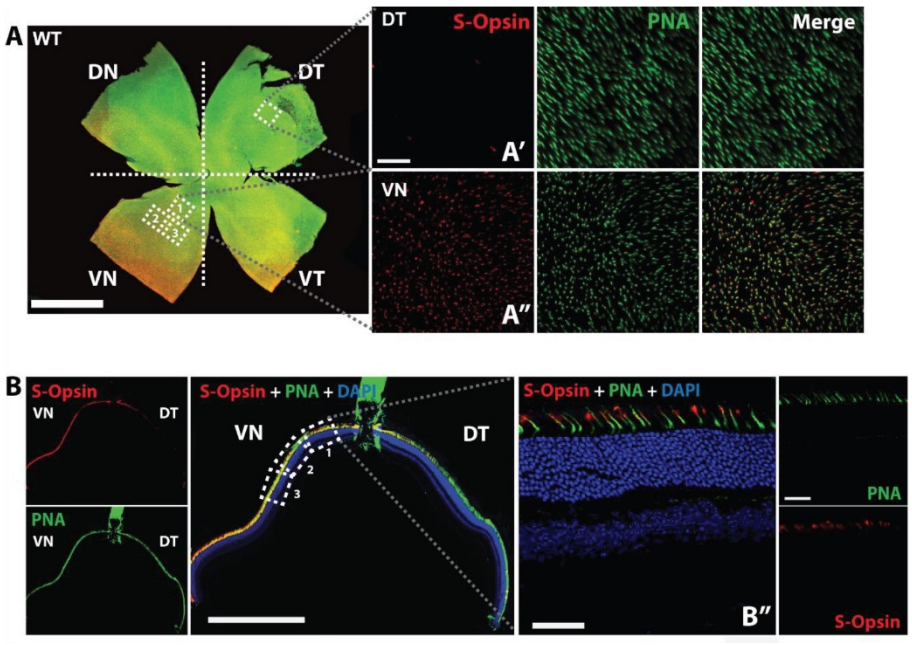


Figure S4. Methodology for the quantification of photoreceptors and other retinal cells types

(A) Representative mosaic reconstruction from 10x images from S-Opsin/PNA double-stained retinal flat mounts from P30 WT mice. (A', A'') 60x confocal microscopy images from retinal flat mount of the DT (A') and from the VN (A'') portion indicate with dashed rectangles in (A). (A') Note the almost complete absence of S-cone (S-Opsin+ cells) in the DT portion where almost only M-pure cones are present. (A") Note that in the VN portion approximately all cones express S-Opsin. The majority of these are dual-photoreceptors expressing both S and M Opsin and a minority of pure S-cone (data not shown). For PRs quantification, 3 images from the VN portion were quantified and averaged as shown by the dashed rectangles in (A). (B) 20x Mosaic Image of the whole retinal section at the level of the optic nerve of an orientated block. The eyes are oriented in the blocks in order to always have VN and DT in opposite positions with respect to the optic nerve on sections. For quantification on retinal sections, 3 consecutive images for sample (as shown by the boxes 1, 2, 3) were taken in the VN side, then quantified and averaged. (B') Representative 60x confocal microscopy image from retinal section used for quantification. Ventro-nasal (VN), dorso-temporal (TD). *n*= 3 to 6 retinas were used for genotype. Scale bars: (A) 1 mm, (B) 0,5mm.

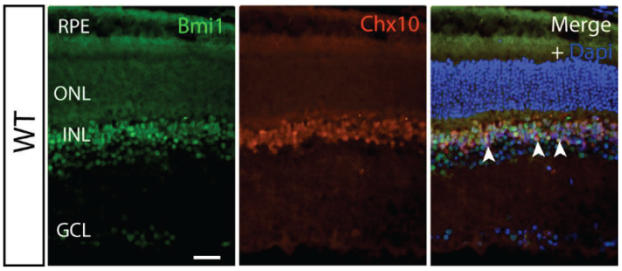


Figure S5. Bmi1 is highly expressed in bipolar neurons

IF analysis of a WT mouse retina (cryosections) using Chx10 and Bmi1 antibodies. Note Bmi1 expression in Chx10-positive cells of the INL. Retinal pigment epithelium (RPE); outer nuclear layer (ONL); inner nuclear layer (INL); ganglion cell layer (GCL). Scale bar: 20µm.

Figure S6. *Chk2* deletion partially improves the *Bmi1*$^{-/-}$ phenotype

Bmi1$^{-/-}$ */Chk2*$^{-/-}$ and *Chk2*$^{+/-}$ mice at P30. Note: *Bmi1*$^{-/-}$ */Chk2*$^{-/-}$ mice were healthier than *Bmi1*$^{-/-}$ mice but remained smaller than *Chk2*$^{+/-}$ control mice.

Table S1. Statistical analysis of ERG parameters

Global values from ERG experiment. Each parameter is analyzed by repeated measures ANOVA. The comparisons between the ERG parameters values at the Vmax are analyzed by one-way ANOVA. All values are mean ± SEM. N.S. = not significant; (*) P ≤ 0.05; (**) ≤ 0.01; (***) ≤ 0.001;

	Photopic system			Scotopic system		
	WT	+/-	-/-	WT	+/-	-/-
Overall a-wave amplitude	F(2,7) = 6.195, p = 0.028 *			F(2,7) = 1.297, p = 0.332 N.S.		
a-wave amplitude at Vmax	-215.4 ± 32.69 µV	-182.5 ± 15.29 µV	-88.35 ± 6.706 µV	-116.1 ± 12.17 µV	-117.4 ± 12.22 µV	-138.0 ± 10.05 µV
Comparison with the other groups (p-value)	+/- = 0.375 N.S.	WT = 0.375 N.S.	WT = 0.008 **	+/- = 0.939 N.S.	WT = 0.939 N.S.	WT = 0.226 N.S.
	-/- = 0.008 **	-/- = 0.039 *	+/- = 0.039 *	-/- = 0.226 N.S.	-/- = 0.281 N.S.	+/- = 0.281 N.S.
Overall a-wave implicit time	F(2,7) = 5.127, p = 0.043 *			F(2,7) = 0.596, p = 0.577 N.S.		
a-wave implicit time at Vmax	8.000 ± 0.4082 ms	8.667 ± 0.3333 ms	9.333 ± 0.3333 ms	21.75 ± 2.175 ms	19.67 ± 0.3333 ms	22.00 ± 2.517 ms
Comparison with the other groups (p-value)	+/- = 0.246 N.S.	WT = 0.246 N.S.	WT = 0.039 *	+/- = 0.484 N.S.	WT = 0.484 N.S.	WT = 0.932 N.S.
	-/- = 0.039 *	-/- = 0.275 N.S.	+/- = 0.275 N.S.	-/- = 0.932 N.S.	-/- = 0.464 N.S.	+/- = 0.464 N.S.
Overall b-wave amplitude	F(2,7) = 18.970, p = 0.001 ***			F(2,7) = 10.440, p = 0.008 **		
b-wave amplitude at Vmax	333.8 ± 34.78 µV	241.9 ± 17.32 µV	132.3 ± 9.479 µV	509.5 ± 24.04 µV	603.7 ± 51.01 µV	371.4 ± 40.33 µV
Comparison with the other groups (p-value)	+/- = 0.044 *	WT = 0.044 *	WT = 0.001 ***	+/- = 0.112 N.S.	WT = 0.112 N.S.	WT = 0.032 *
	-/- = 0.001 ***	-/- = 0.029 *	+/- = 0.029 *	-/- = 0.032 *	-/- = 0.004 **	+/- = 0.004 **
Overall b-wave implicit time	F(2,7) = 20.852, p = 0.001 ***			F(2,7) = 0.034, p = 0.966 N.S.		
b-wave implicit time at Vmax	27.25 ± 0.6292 ms	31.00 ± 1.000 ms	39.33 ± 2.848 ms	57.00 ± 1.472 ms	56.67 ± 1.453 ms	53.00 ± 2.517 ms
Comparison with the other groups (p-value)	+/- = 0.136 N.S.	WT = 0.136 N.S.	WT = 0.001 ***	+/- = 0.899 N.S.	WT = 0.899 N.S.	WT = 0.158 N.S.
	-/- = 0.001 ***	-/- = 0.010 **	+/- = 0.010 **	-/- = 0.158 N.S.	-/- = 0.217 N.S.	+/- = 0.217 N.S.

9 Annex IV: Off-target effect of the BMI1 inhibitor PTC596 drives epithelial-mesenchymal transition in glioblastoma multiforme

This third annex is an article, on which I have worked during my Ph.D., and was published on Precision Oncology in 2020. In this article we continued the work on the functions of BMI1, specifically its role in maintaining the stemness of cancer stem cells.

npj | Precision Oncology

www.nature.com/npjprecisiononcology

ARTICLE OPEN

Off-target effect of the BMI1 inhibitor PTC596 drives epithelial-mesenchymal transition in glioblastoma multiforme

Anthony Flamier[1,3,4], Mohamed Abdouh[1,4], Rimi Hamam[1], Andrea Barabino[1], Niraj Patel[1], Andy Gao[1], Roy Hanna[1] and Gilbert Bernier[1,2*]

Glioblastoma multiforme (GBM) is an incurable primary brain tumor containing a sub-population of cancer stem cells (CSCs). Polycomb Repressive Complex (PRC) proteins BMI1 and EZH2 are enriched in CSCs, promoting clonogenic growth and resistance to genotoxic therapies. We report here that when used at appropriate concentrations, pharmaceutical inhibitors of BMI1 could efficiently prevent GBM colony growth and CSC self-renewal in vitro and significantly extend lifespan in terminally ill tumor-bearing mice. Notably, molecular analyses revealed that the commonly used PTC596 molecule targeted both BMI1 and EZH2, possibly providing beneficial therapeutic effects in some contexts. On the other hand, treatment with PTC596 resulted in instant reactivation of EZH2 target genes and induction of a molecular program of epithelial–mesenchymal transition (EMT), possibly explaining the modified phenotype of some PTC596-treated tumors. Treatment with a related but more specific BMI1 inhibitor resulted in tumor regression and maintenance of cell identity. We conclude that inhibition of BMI1 alone is efficient at inducing GBM regression, and that dual inhibition of BMI1 and EZH2 using PTC596 may be also beneficial but only in specific contexts.

npj Precision Oncology (2020)4:1 ; https://doi.org/10.1038/s41698-019-0106-1

INTRODUCTION

Glioblastoma multiforme (GBM) represents the most common brain malignancy in adults. However, current treatments are mostly effective at reducing intracranial brain pressure and the alkylating agent Temozolide (Temodal) can increase lifespan by ~4 months. The median lifespan of patients at the time of diagnosis is still 9–12 months. An effective treatment is thus critically needed.[1–3] GBMs are highly heterogeneous tumors containing a relatively rare sub-population of cancer-initiating cells expressing the CD133 (*PROM1*) cell surface antigen.[4–6] Based on cell culture and xenotransplantation experiments, it was shown that CD133+ cells behave like neural stem cells, express stem cell markers, and are able to generate new brain tumors in serial transplantations.[5–7] The CD133+ cancer stem cell (CSC) fraction also represents the radio-resistant cell population in GBM and is believed to be responsible for brain tumor reoccurrence after radiotherapy treatments.[8–10] Importantly, cell lines grown under serum conditions are not representative of the phenotype of primary GBM tumors.[7]

Transcription factors (TFs), such as SOX2 and OLIG2, have been reported as key molecular cues for gliomagenesis and tumor maintenance.[4,11,12] More specifically, TLX, ZFHX4, and MLL5 are TFs and chromatin remodelers overexpressed in glioma CSC population and important for their self-renewal.[13–15] Although the interconnection between these factors remains unclear, we can hypothesize that CSC identity relies on a self-sustaining network of TFs. Notably, it has been proposed that the CSC phenotype and other key features of cancer cells may be driven by an epigenetic circuitry in conjuncture with genetic mutations.[16]

Polycomb Repressive Complexes (PRCs) form large multimeric complexes responsible for the remodeling of the chromatin and can promote gene silencing through specific histone modifications.[17] They are classically subdivided into two groups, namely PRC1 (which includes BMI1, RING1a, and RING1b/RNF2) and PRC2 (which includes EZH2, EED, and SUV12).[18] The subsequent histone modifications induced by PRC1 and PRC2 complexes are critical to maintain stable silencing of both euchromatin and facultative heterochromatin.[19–21] The PRC2 is able to silence chromatin through its histone H3 tri-methylase activity at lysine 27 ($H3K27^{me3}$) while the PRC1 uses histone H2A mono-ubiquitin ligase activity at lysine 119 ($H2A^{ub}$).[19–21] Studies on the course of mouse development showed that $H3K27^{me3}$ deposition by PRC2 is thought to be a pioneer event required for PRC1 recruitment at developmental genes. Conversely, $H2A^{ub}$ mark by PRC1 may be necessary for $H3K27^{me3}$ deposition and maintenance in somatic cells, constituting a positive feedback loop.[22] Members of the PRC have been identified as proto-oncogenes in human cancers.[23–28] *BMI1* (B-cell specific Moloney murine leukemia virus integration site 1) is one of those and initially described as an oncogenic partner in lymphomagenesis. *BMI1* has been found to be overexpressed in several cancers and been shown to be crucial for cancer cell survival in medulloblastoma and glioblastoma.[10,29–36] Consequently, *BMI1* inhibition in human or mouse GBM cells results in impaired CSC self-renewal and absence of tumor formation in grated mice, and this independently of a functional *Ink4a* locus.[29,36] Intriguingly, *BMI1* overexpression can confer self-renewal properties and is apparently sufficient to "reprogram" mouse astrocytes into neural stem cells or mouse retinal progenitors into retinal "stem cells".[37,38]

GBM tumors have been classified into three major sub-types based on differential gene expression. The *proneural/neural* subtype enriched for *DLL3*, *OLIG2*, *ASCL1*, *PDGFRA*, *IDH1*, and *PROM1*; the *classical* subtype enriched for *FOXO3*, *NES* (Nestin), *EGFR* and *AKT2*; and the *mesenchymal* subtype enriched for *CD44* (SSEA1), *CHI3L1*, *NF1*, *TIMP1*, and *TGFβ*.[39] Notably, it was recently suggested that *EZH2* is enriched in the proneural subtype and

[1]Stem Cell and Developmental Biology Laboratory, Hôpital Maisonneuve-Rosemont, 5415 Boul. l'Assomption, Montréal H1T 2M4, Canada. [2]Department of Neurosciences, University of Montreal, Montreal, Canada; [3]Present address: Whitehead Institute of Biomedical Research, 455 Main Street, Cambridge 02142 MA, USA. [4]These authors contributed equally: Anthony Flamier, Mohamed Abdouh. *email: gbernier.hmr@ssss.gouv.qc.ca

Published in partnership with The Hormel Institute, University of Minnesota

npj nature partner journals

The Hormel Institute University of Minnesota

BMI1 in the mesenchymal subtype.[40] Hence, proneural tumors were apparently resistant to BMI1 inhibition using PTC596, and mesenchymal tumors apparently resistant to EZH2 inhibition using two distinct inhibitors. Based on this, it was proposed that dual inhibition of BMI1 and EZH2 is more efficient at eradicating GBM then when using BMI1 or EZH2 inhibitors alone.[40]

Herein, we present evidences that independently of BMI1 expression level or of the GBM subtype, GBM neural spheres are sensitive to the related BMI1 inhibitors PTC596 and A1016 at a range between 5 and 50 nM. We found that in contrast with A1016, PTC596 markedly interfered with EZH2, FOXG1, and SOX2 protein levels. RNA sequencing (RNA-seq) analyses confirmed that A1016 more closely aligned than PTC596 with the *BMI1* knockout gene expression profile, and that PTC596-treated GBM spheres showed activation of an epithelial to mesenchymal transition (EMT) molecular program and de-repression of PRC2-target genes. In terminally ill mice bearing intracranial tumors, treatments with high concentrations of PTC596 significantly extended median and maximal lifespan. In some but not all grafts, however, relapsing PTC596-treated tumors showed reduced BMI1, EZH2, and SOX2 expression, suggesting epigenetic drift. We conclude that inhibition of BMI1 is highly efficient at eliminating GBM tumors and that treatments with PTC596, which targets BMI1 and EZH2, may be only beneficial in specific contexts.

RESULTS

PTC596 is a cell-permeable small compound capable of inducing BMI1 protein proteosomal degradation at nanomolar concentrations.[41] The compounds can cross the blood–brain barrier and is orally administrable.[40] To evaluate its utility for brain cancer treatment, we exposed the patient-derived GBM0811 cell line maintained and grown as neurospheres to various concentrations of PTC596, or to A1016, a closely related molecule. After 7 days of treatment, colony growth and cell viability were measured, revealing that both drugs efficiently reduced colony growth starting at 5 nM (Fig. 1a). This also correlated with an important reduction of cell viability in the remaining spheres (Fig. 1b). To evaluate the effect on CSC self-renewal, cancerous neurospheres of the GBM1205 cell line were treated for 7 days with 5 nM of drugs, washed, and maintained for an additional 25 days in drug-free media. Remaining spheres were then dissociated at 4600 viable cells/well in fresh media and maintained for an extra 18 days. We found that both drugs dramatically affected the number and size of newly formed colonies, suggesting depletion of the CSC population (Fig. 1c, d). Likewise, neurospheres of the GBM0410 cell line exposed to 5 nM of drug for 2 days were dissociated in drug-free media (Fig. 1e). After 2 weeks, newly formed neurospheres were re-dissociated in drug-free media to measure secondary sphere formation (Fig. 1e). This revealed that acute exposure of the GBM spheres was sufficient to affect the formation of primary and secondary neurospheres, again suggesting depletion of the CSC population (Fig. 1f).

To test the effect on colony growth, GBM cells were cultured as a monolayer on matrigel in neural stem cell media until large colonies were present. The cultures were then exposed to 5 nM of BMI1 inhibitors for 3 days (Fig. S1A). In all cell lines tested, we observed a severe reduction in colony size and evidences of cell death (Fig. S1A–C). Moreover, we noticed that viable GBM cells rapidly underwent neural (MAP2) and glial (GFAP) differentiation upon exposure to PTC596, as revealed using immunofluorescence (Fig. S1D). Real-time RT-PCR analyses also revealed that *BMI1* mRNA expression was not affected by the drug treatment, but that *PROM1*, a marker of CSC, was decreased in a dose-dependent manner upon treatment with both inhibitors (Fig. S1E). Taken together, these results were consistent with our previous work showing that *BMI1* knockdown resulted in GBM stem cell differentiation and in the loss of self-renewal capacity.[29]

To study the molecular effects of BMI1 inhibitors, we performed western blot analyses on GBM neurospheres treated for 24 h with increasing concentrations of PTC596 or A1016. DMSO was used as a control. BMI1 inhibition in the GBM0811 and GBM1205 cell lines was observed already at 5 nM, with maximum inhibition reached between 5 and 50 nM, depending on the cell line (Fig. 2a, b). In time course studies, BMI1 inhibition was observed after ~5 h with A1016, and after ~8 h with PTC596 (not shown). To test the effect on the CSC phenotype, we measured CD133, EZH2, FOXG1, and SOX2 protein expression using immunoblot.[29,42–46] In both cell lines tested, BMI1 inhibition for 24 h resulted in a significant reduction of CD133 (Fig. 2b, c) and $H2A^{ub}$ (Fig. 3a) levels. Surprisingly, we found that the expression of EZH2, FOXG1, and SOX2 was also significantly reduced in PTC596-treated cells, whereas it was almost unchanged in A1016-treated cells (Fig. 2b–d). Using immunofluorescence on GBM sphere sections, we observed that Nestin expression remained unaffected after treatment with PTC596 or A1016 for 24 h (Fig. 2e). However, while EZH2 expression remained in normal ranges in A1016-treated spheres, it was highly reduced in spheres treated with PTC596 (Fig. 2f).

BMI1 is highly expressed in mature human and mouse cortical neurons and its expression is reduced in neurons from Alzheimer's disease patients.[47,48] Furthermore, acute BMI1 knockdown in cultured human cortical neurons leads to severe neurodegeneration,[48] thus raising substantial concerns about possible side effects of BMI1 inhibitors on normal brain function. To investigate this, we generated day in vitro 35 post-mitotic cortical neurons through directed differentiation of human embryonic stem cells or induced pluripotent stem cells.[48] Cultured neurons were treated with 100 nM of PTC596 or A1016 for 24 h and analyzed by immunoblot. Notably, drug-treated neurons were apparently healthy and presented elevated levels of BMI1 and $H2A^{ub}$, suggesting a concomitant increase in the biochemical activity of the PRC1 (Fig. S2A). Similarly, oral administration of PTC596 or A1016 at 12 mg/kg in post-natal day 45 mice was not accompanied by Bmi1 down-regulation when whole cortices were analyzed by immunoblot (Fig. S2B) or immuno-histochemistry (not shown) 24 h after treatment. To support these observations, stem cell-derived post-mitotic neurons and neural progenitors were exposed to PTC596 or A1016 for 24 h and test for cell viability. While neurons exposed to BMI1 inhibitors were relatively unaffected, neural progenitors showed decreased viability as revealed using the MTT assay or by measuring the proportion of dying cells (Fig. S3A, B). To evaluate the mechanism of action of the new A1016 molecule on BMI1, lysates from neural progenitors exposed to 100 nM of A1016 for 24 h were immuno-precipitated with an antibody against poly-ubiquitin at lysine 48, a mark for protein degradation by the proteasome. This revealed increased poly-ubiquitinated BMI1 in A1016-treated cells, as measured using western blot analysis (Fig. S3C). These results suggested that the inhibitory activity of PTC596 and A1016 on BMI1 is cell cycle dependent and thus less likely to affect the viability of post-mitotic neurons.

EZH2 is the catalytic sub-unit of the PRC2 and is required for $H3K27^{me3}$ deposition and maintenance.[49] On the other hand, the BMI1//RING1 complex regulates $H2A^{ub}$ deposition, possibly also directly or indirectly affecting $H3K9^{me3}$ levels in normal somatic cells.[50] Consistently, we found that treatment of GBM spheres (GBM1205) with PTC596 for 24 h resulted in robust down-regulation of $H2A^{ub}$ levels (Fig. 3a). Notably, this was also accompanied by down-regulation of $H3K27^{me3}$ and $H3K9^{me3}$ levels (Fig. 3a). The heterochromatin is tightly associated with the nuclear envelope and loss of $H3K9^{me3}$ can result in nuclear lamina disintegration.[51] Using immunofluorescence, we observed that GBM spheres (GBM1909) treated for 24 h with PTC596 showed reduced $H3K9^{me3}$ levels and disintegration of the nuclear lamina (Fig. 3b). This revealed that PTC596 treatment in GBM cells

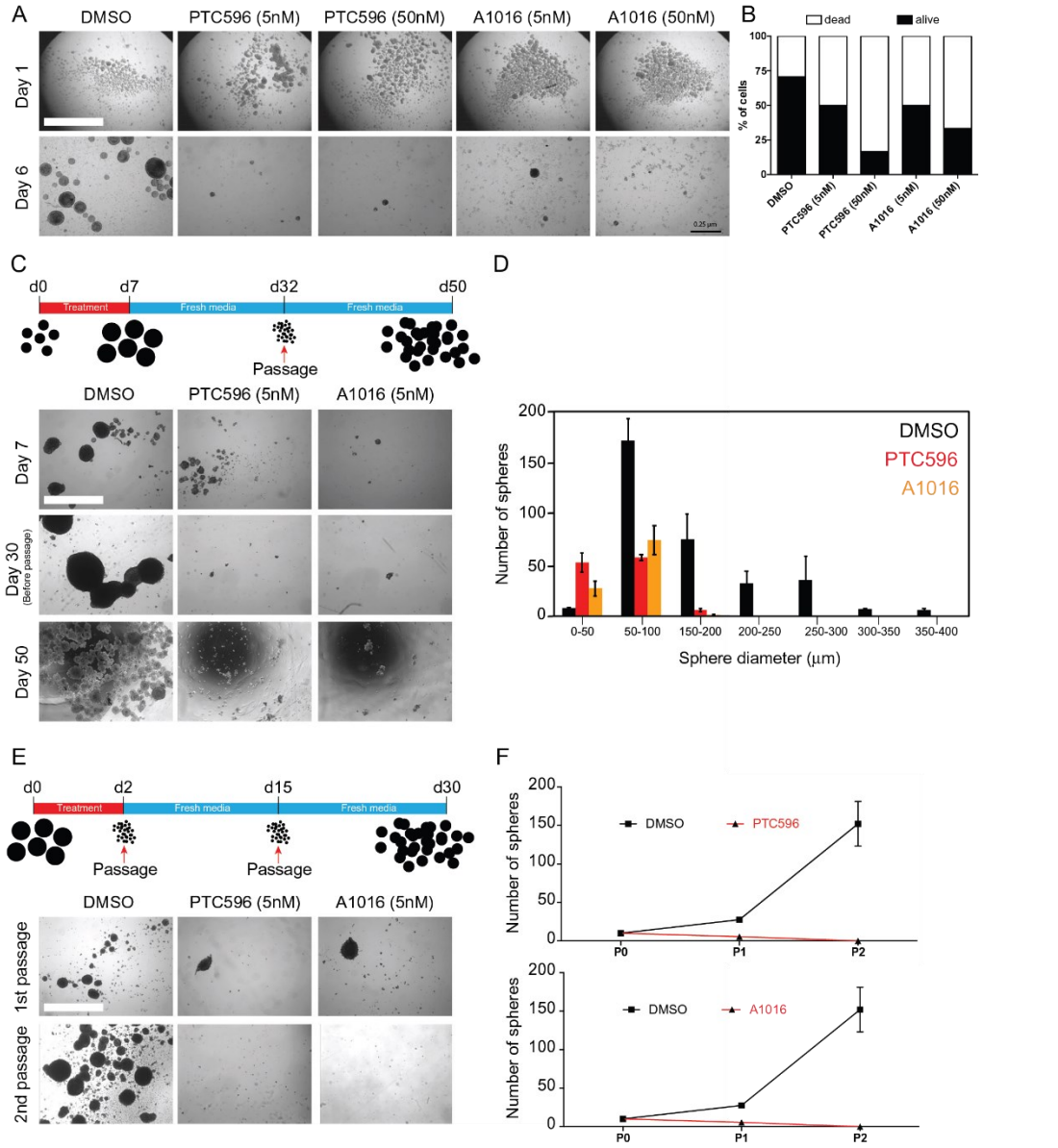


Fig. 1 BMI1 inhibitors impair GBM colony growth and cancer stem cell self-renewal. a Representative images of GBM0811 cells treated for 6 days by DMSO, PTC596 (5 and 50 nM), and A1016 (5 and 50 nM). Scale bar: 2.5 mm. **b** Quantification of cell viability in GBM0811 cells treated for 6 days by DMSO, PTC596 (5 and 50 nM) and A1016 (5 and 50 nM). **c** GBM1205 cell growth assay upon acute treatment (7 days) with BMI1 inhibitors PTC596 and A1016. Top: scheme of the assay. Bottom: representative images at each time point. Scale bar: 2.5 mm. **d** Size distribution of the spheres at day 50 from experiment in **c** after PTC596 or A1016 treatments in comparison to DMSO-treated cells. **e** Colony-forming assay after acute treatment (2 days) with BMI1 inhibitors and two passages of the GBM0410 cells. Top: scheme of the assay. Bottom: representative images at each time point. Scale bar: 2.5 mm. **f** Quantification of the number of spheres after one and two passages in PTC596 (Top) or A1016 (Bottom) treated cells.

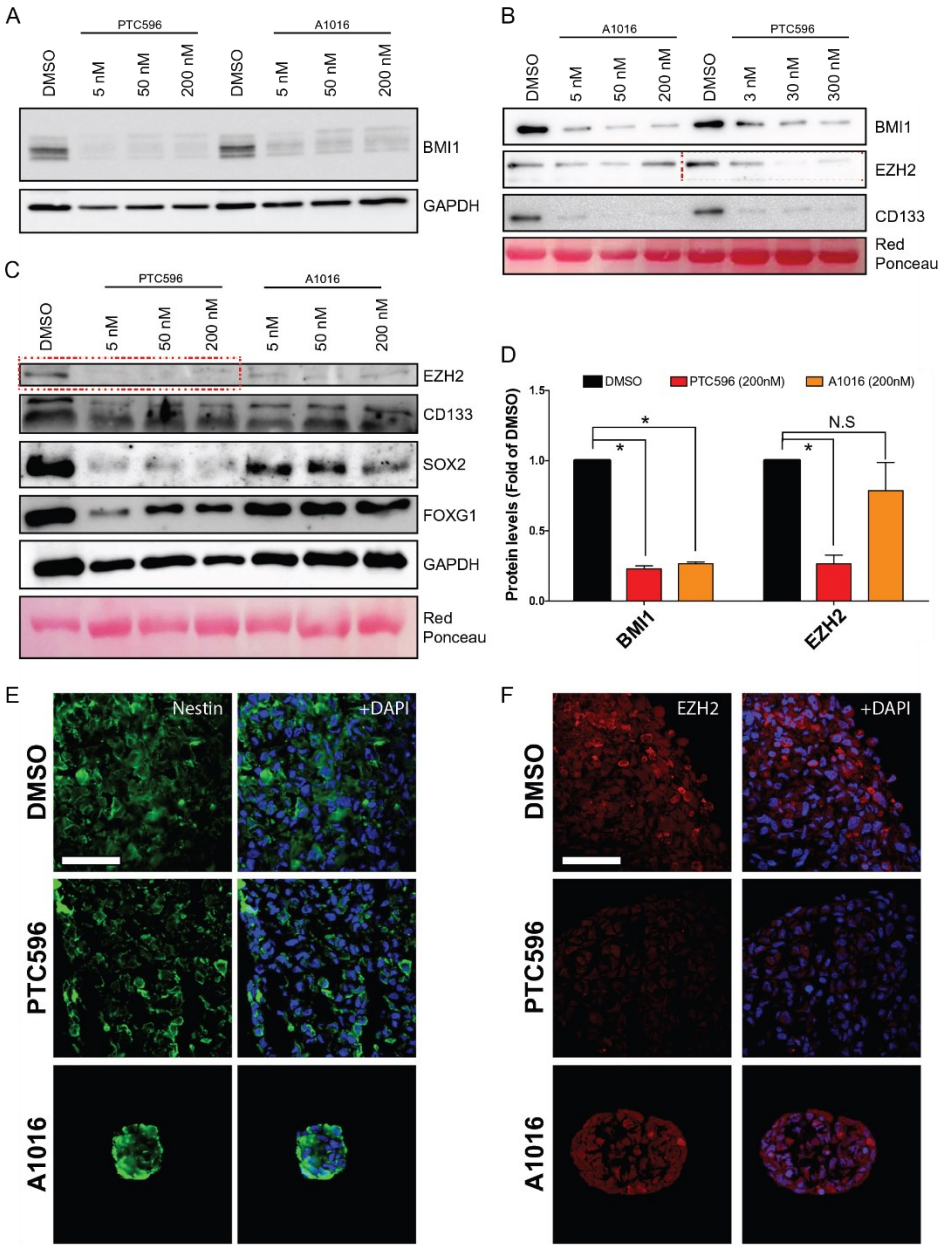


Fig. 2 PTC596 inhibits BMI1 and EZH2 in GBM neurospheres. a Immunoblot of BMI1 in GBM line 1205 after treatment with increasing concentrations of BMI1 inhibitors. GAPDH is used as a loading control. **b** Immunoblot of BMI1, EZH2, and CD133 in GBM line 0811 after treatment with increasing concentrations of BMI1 inhibitors. Actin and Red ponceau are used as a loading control. Note the reduced EZH2 levels in PTC596-treated cells (dashed red box). **c** Immunoblot of SOX2, EZH2, FOXG1, and CD133 in GBM line 1205 after treatment with increasing concentrations of BMI1 inhibitors. GAPDH and Red ponceau are used as a loading control. Note the reduced EZH2 levels in PTC596-treated cells (dashed red box). **d** Quantification of BMI1 and EZH2 protein levels in GBM cells ($n = 2$ cell lines) after treatment with BMI1 inhibitors (200 nM). All values are mean ± SEM. *P value <0.05. **e** Immunofluorescence for Nestin on GBM line 1205 neurospheres cryo-sectioned after 48 h of treatment with DMSO, PTC596, or A1016. Scale bar: 20 μm. **f** Immunofluorescence for EZH2 on GBM line 1205 neurospheres cryo-sectioned after 48 h of treatment with DMSO, PTC596, or A1016. Scale bar: 20 μm.

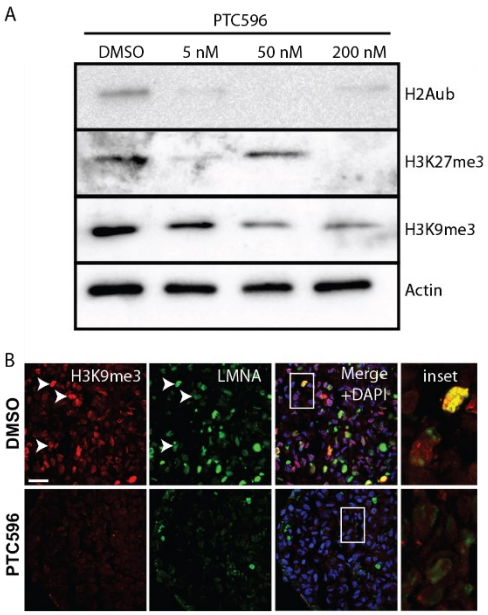


Fig. 3 PTC596 treatment perturbs several histone modifications and the integrity of the nuclear lamina. a Immunoblot analysis of GBM neurospheres (GBM1205 cell line) treated with PTC596 for 24 h. Note the dose-dependent reduction in $H2A^{ub}$, $H3K27^{me3}$, and $H3K9^{me3}$ levels. **b** Immunofluorescence analysis on sections of GBM neurospheres (GBM1909 cell line) treated with PTC596 for 24 h. Note the reduced $H3K9^{me3}$ land LaminA/C (LMNA) levels in drug-treated GBM spheres. Scale bar: 20 µm.

significantly perturbed PRC1- and PRC2-mediated histone modifications and structure of the nuclear envelope.

Next, we further analyzed the novel patient-derived cell line, GBM1205, using RNA-seq. When compared to primary GBM tumors from the IVY GAP repository, we determined that the GBM1205 cell line had a mixed phenotype, showing a proneural phenotype with some classical features (Fig. 4a). The cell line was also able to generate hemorrhagic brain tumor in grafted immune-deficient mice (Fig. 4b). To dissect the pioneer molecular changes resulting from *BMI1* inactivation, we targeted *BMI1* exon 1 using CRISPR/Cas9 technology to generate *BMI1* knockout ($BMI1^{KO}$) GBM cells.[48] A non-targeting template guide RNA was used as a negative control. Control ($BMI1^{+/+}$) and $BMI1^{KO}$ cells were collected 24 h post-transfection for western blot and RNA-seq analyses. We estimated that ~70% of the cells were targeted, and $BMI1^{KO}$ cell extracts exhibited reduced BMI1 and $H2A^{ub}$ levels, confirming *BMI1* deletion (Fig. 4c). Differential expression analysis revealed that most de-regulated genes in $BMI1^{KO}$ cells were upregulated. This is consistent with BMI1, working within the PRC1, functioning as a gene-silencing factor (Fig. 4d).[21,52,53] Gene ontology (GO) annotation of the most upregulated genes in $BMI1^{KO}$ showed significant appearance of interferon, cell metabolism, chromatin remodeling, and apoptosis-related GO terms (Figs. 4e and S4). Using Gene Set Enrichment Analysis (GSEA), we found an enrichment for apoptosis gene set in $BMI1^{KO}$, but gene expression levels remained relatively low (Fig. 4f). Notably, $BMI1^{KO}$ GBM cells were also enriched in genes regulated by NF-κB in response to TNFα signaling, suggesting an inflammation-related cellular response (Fig. 4f). Notably, the CSC gene expression profile was also altered upon *BMI1* inactivation, including that of *PROM1*. Hence, the expression level of 14 stem cell-related genes, all associated with a proneural GBM phenotype, was significantly reduced in $BMI1^{KO}$ cells, suggesting that BMI1 may be involved in the maintenance of the proneural CSC phenotype in GBM (Figs. 4g and S5).[54,55]

To assess the specificity of the compounds, we compared the RNA-seq profiles of DMSO-, PTC596-, and A1016-treated cells with that of $BMI1^{KO}$ cells (Figs. 5 and S6). When compared to DMSO, PTC596-treated cells showed more de-regulated genes than A1016-treated cells, many of which being downregulated (Fig. 5a). In contrast, most de-regulated genes in A1016-treated cells were upregulated (Fig. 5a), similarly as observed in $BMI1^{KO}$ cells (Fig. 4d). When directly compared to $BMI1^{KO}$ cells, PTC596-treated cells presented 4099 significantly de-regulated genes compared to the 2732 de-regulated genes observed in A1016-treated cells (Fig. 5b). Taken together, this suggested significant off-target effects of the PTC596 compound. Using GSEA, we found that genes upregulated in response to *BMI1* knockdown in cancer cells were also upregulated in PTC596 and A1016-treated GBM cells, suggesting efficient drug-mediated BMI1 inhibition (Fig. S6A, B). Treatment with both compounds also succeeded at reducing the expression of glioblastoma proneural genes as observed in $BMI1^{KO}$ cells (Fig. 5c). In contrast to $BMI1^{KO}$ or A1016-treated cells however, PTC596 treatment resulted in a molecular signature resembling a mesenchymal phenotype (Fig. S6C), and an enrichment for a gene set involved in EMT (Fig. 5c). Likewise, SOX2-responsive genes were significantly downregulated in PTC596-treated cells, but not in $BMI1^{KO}$ or A1016-treated cells (Fig. 5c). Finally, GSEA revealed a significant up-regulation of EZH2 and PRC2-repressed genes in PTC596-treated cells using three independent gene sets (Figs. 5c and S6D).

To evaluate the efficiency of PTC596 at eliminating GBM in vivo, we injected the GBM1205 cell line maintained as neural stem cells in serum-free media in the lateral ventricle of the cerebral cortex of $NOD/SCID^{IL-2}$ mice ($n = 10$ mice). PTC596 was given orally at 12 mg/kg every 3 days for 25 days, and this starting after the first death in the cohort. This protocol was thus designed as an end-stage disease treatment. When compared to HPMC-treated mice (control), PTC596-treated animals showed an extension of median lifespan of 26 days, with one treated animal that was tumor-free 3 months post-treatment (Fig. 6a). Likewise, mice grafted with the GBM0811 cell line ($n = 16$ mice) and treated with HPMC or PTC596 at 6 mg/kg every 3 days also showed an extension of median lifespan by 17 days (Fig. S7). Notably, we could confirm BMI1 and EZH2 down-regulation in the brain tumor during the treatment period and their reactivation 15 days post-treatment in an animal showing relapse (Fig. S8). To simultaneously compare the efficiency of both inhibitors, we injected the GBM0811 cell line in 25 mice. The cohort was composed of 5 HPMC-treated mice, 10 PTC596-treated mice (6 and 12 mg/kg, every 3 days), and 10 A1016-treated mice (6 and 12 mg/kg, daily). When compared to untreated mice, we found that PTC596 and A1016 given at 6 mg/kg extended median survival by 6 and 25 days, respectively (Fig. 6b). At 12 mg/kg, PTC596 extended median lifespan by 41 days, with one tumor-free animal 3 months post-treatment. Mice treated daily with A1016 at 12 mg/kg succumbed from drug toxicity few days after the end of the treatment. Notably, however, all animals treated with A1016, but one, were tumor-free (Fig. 6b). When compared to the tumor of control mice, the tumor of PTC596-treated mice with relapse (mice #5 and #6) had lost many the original characteristics, including reduced BMI1, EZH2, and SOX2 expression (Figs. 6c and S9). In contrast, in the one animal treated with A1016 and that still presented a small tumor of 0.6 mm in diameter, the tumor remained BMI1 and SOX2 positive, although EZH2 levels were reduced (Figs. 6d and S9). These results suggested significant therapeutic effects of both BMI1 inhibitors when used as single agents for the treatment of GBM.

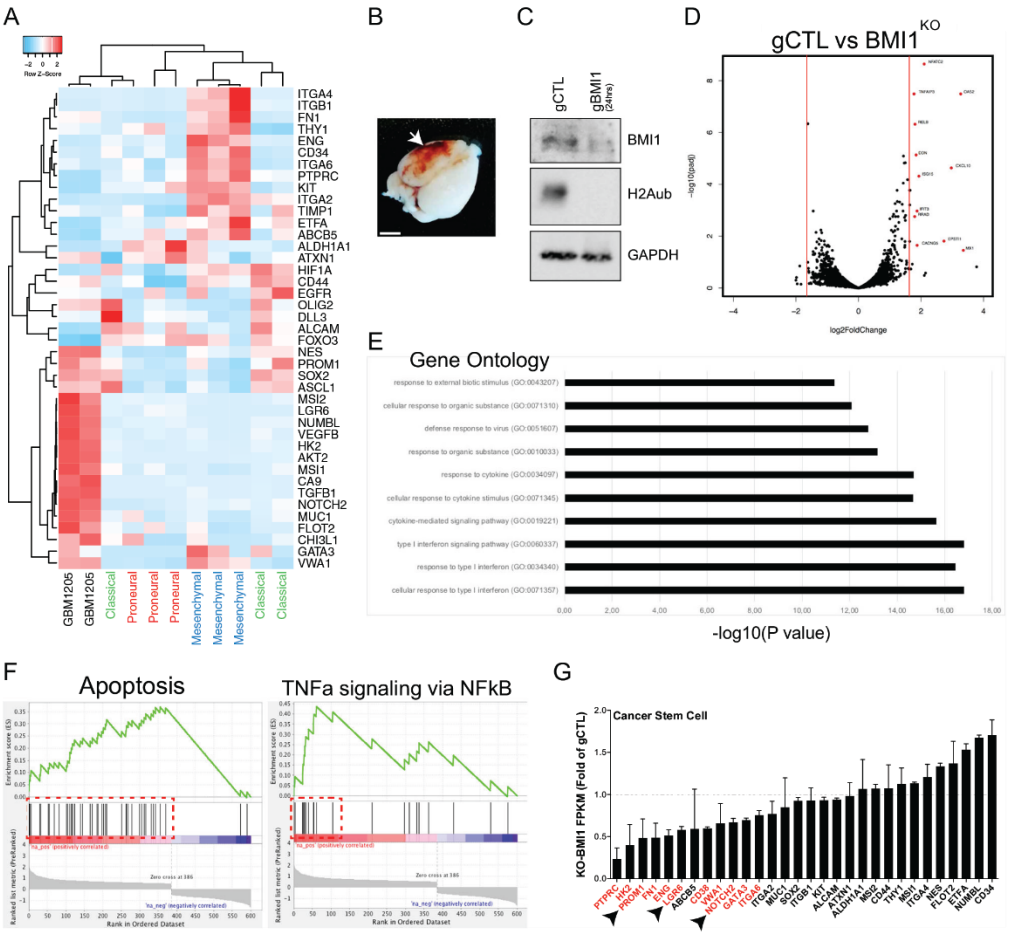


Fig. 4 Acute *BMI1* knockout in GBM induces an inflammation-related response and down-regulation of cancer stem cell genes. **a** Heatmap of gene expression for glioblastoma genes in GBM line 1205 ($n = 2$) and GBM tumors having either a classical (green), mesenchymal (blue), or proneural (red) phenotype ($n = 3$ for each group). Row *z*-scores are a function of FPKM value. **b** Image of a NOD/SCID mice brain after intracranial injection of the GBM line 1205. Scale bar: 4 mm. **c** Immunoblot of BMI1 and H2Aub in GBM line 1205 transfected with a non-template control guide RNA (gCTL) or a guide RNA targeting BMI1 (gBMI1) with a plasmid overexpressing Caspase 9. GAPDH is used as a loading control. **d** Volcano plot showing the most dysregulated genes in BMI1^{KO} versus gCTL GBM line 1205. Red dots show significantly dysregulated genes. **e** Selection of Gene Ontology terms for upregulated genes in BMI1^{KO} versus gCTL GBM line 1205. **f** Gene Set Enrichment Analysis (GSEA) showing enrichment for apoptosis and TNFa gene sets in BMI1^{KO} versus gCTL GBM line 1205. **g** Gene expression levels of cancer stem cells genes in BMI1^{KO} versus gCTL GBM line 1205. Red annotations show significantly downregulated genes.

DISCUSSION

Epigenetic drug therapy may represent a new approach to cure cancers. Herein, we showed that BMI1 inhibitors could efficiently block GBM neural sphere growth and CSC viability in four independent patient-derived GBM cell lines, suggesting that drug-mediated BMI1 inhibition is a versatile treatment for most patients with GBM. Molecular studies revealed that the compounds were able to reduce the BMI1 protein level when used at nanomolar concentration in GBM cell lines having proneural and/or classical phenotypes. When compared to A1016 or *BMI1* knockout, the PTC596 molecule presented additional effects, including inhibition of EZH2, SOX2, and FOXG1. Terminally ill mice bearing brain tumors and treated with PTC596 showed highly extended lifespan. However, analysis of some PTC596-treated animals with relapse revealed that the tumors presented a modified phenotype characterized by reduced BMI1, EZH2, and SOX2 levels.

In a similar strategy to the reprogramming of somatic cells into pluripotent stem cells using a set of four master TFs,[56] the simultaneous overexpression of *SOX2*, *OLIG2*, *POU3F2*, and *SALL2* reprograms differentiated GBM cells into CSC able to drive tumor formation.[16] Likewise, glioma-initiating cells can be obtained from tumor suppressor-deficient astrocytes through the ectopic expression of *SOX2*, *OLIG2*, and *ZEB1*.[57] EZH2 is the catalytic unit of the PRC2 and is overexpressed in GBM. Because EZH2 inactivation impairs cell growth, it prompted interest as a potential target

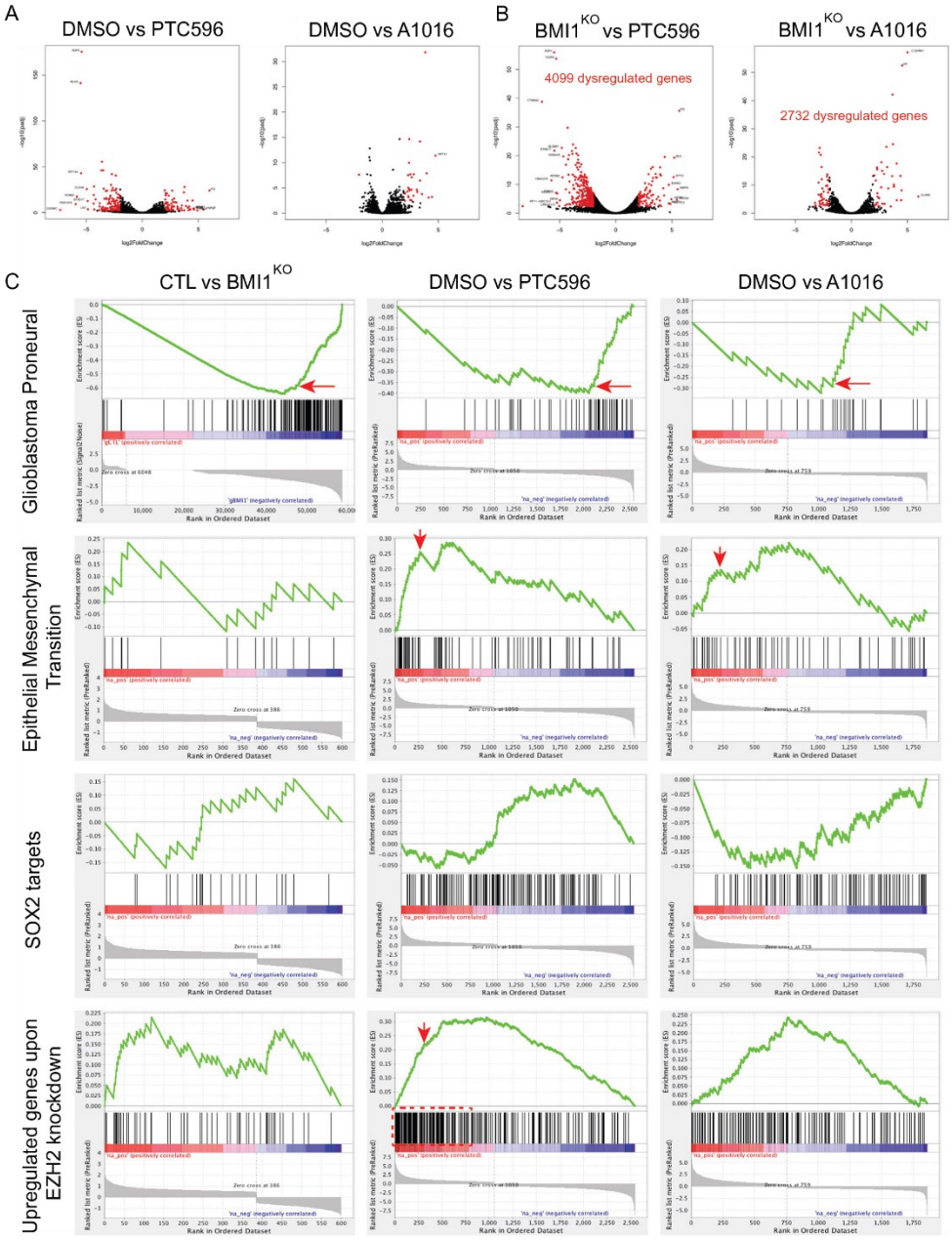
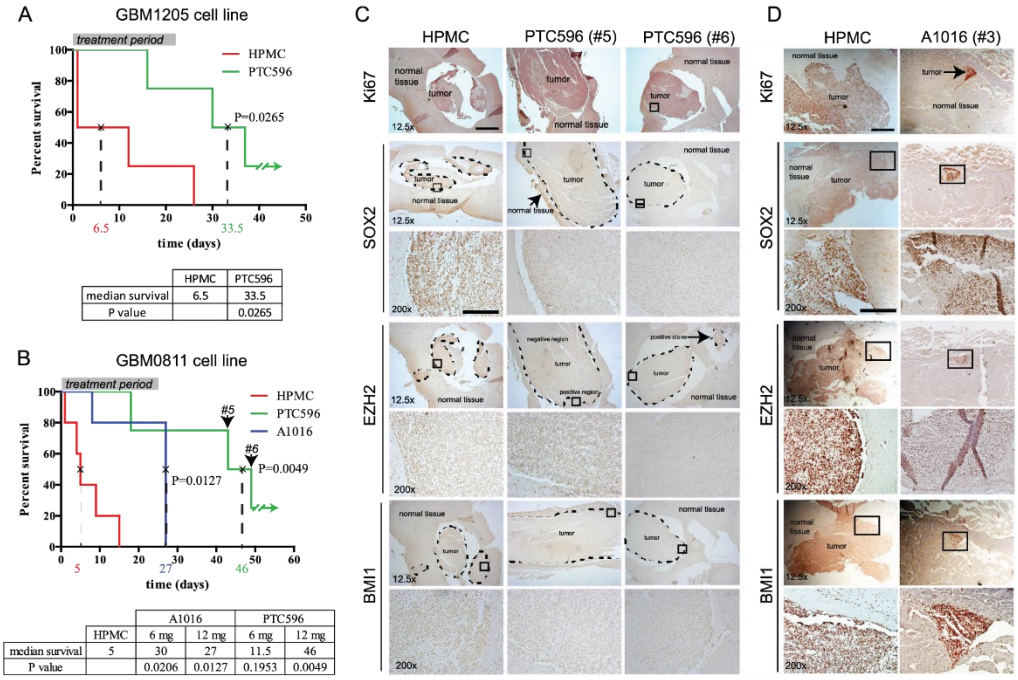


Fig. 5 PTC596 perturbs the expression of EZH2 and SOX2 target genes. a Volcano plot showing the most dysregulated genes in DMSO versus PTC596- or A1016-treated GBM line 1205 for 24 h. Red dots show significantly dysregulated genes. **b** Volcano plot showing the most dysregulated genes in BMI1^{KO} versus PTC596- or A1016-treated GBM line 1205 for 24 h. Red dots show significantly dysregulated genes. **c** Gene Set Enrichment Analysis (GSEA) showing enrichment for glioblastoma proneural, epithelial–mesenchymal transition, SOX2 targets, and EZH2 targets gene sets in BMI1^{KO}, PTC596, or A1016 GBM line 120516. Red arrows and boxes highlight enrichments of interest.



	HPMC	PTC596
median survival	6.5	33.5
P value		0.0265

		A1016		PTC596	
	HPMC	6 mg	12 mg	6 mg	12 mg
median survival	5	30	27	11.5	46
P value		0.0206	0.0127	0.1953	0.0049

Fig. 6 Tumor-bearing mice treated with BMI1 inhibitors show increased lifespan. a Kaplan Meier graph for NOD/SCID mice grafted by GBM line 1205 in the brain and treated with BMI1 inhibitors ($n = 4$ for each group) or vehicle (HPMC; $n = 5$). All values are mean ± SEM. **b** Kaplan Meier graph for NOD/SCID mice grafted by GBM line 0811 in the brain and treated with BMI1 inhibitors ($n = 4$ for each group) or vehicle (HPMC; $n = 5$). All values are mean ± SEM. **c** Immuno-histochemistry for BMI1, SOX2, and EZH2 on brain sections from tumor-bearing mice treated or not (HPMC) with PTC596 (GBM0811 cell line). Ki67 was used to confirm cell proliferation in the tumor. Scale bars 12.5×: 1 mm; 200×: 125 µm. **d** Immuno-histochemistry for BMI1, SOX2, and EZH2 on brain sections from tumor-bearing mice treated or not (HPMC) with A1016 (GBM0811 cell line). Ki67 was used to confirm cell proliferation in the tumor. Scale bars 12.5×: 1 mm; 200×: 125 µm.

against glioma.[29,46,58,59] Although EZH2 has a proto-oncogene function, it can also abrogate tumor transformation. Hence, the inactivation of *Ezh2* impairs mouse GBM tumor growth and extends lifespan. Most notably however, prolonged *Ezh2* inactivation causes a loss of the H3K27^{me3} mark, which induces an activation of some pluripotency markers, resulting to a cell fate change and an aggressive tumor transition.[60] Thus, loss of H3K27^{me3} mark in GBM is predictive for a transition towards a more immature and aggressive phenotype. Notably, Wu et al.[61] have identified recurrent mutations in *H3F3A* and *HIST3H1B* (encoding for histones H3.3 and H3.1) at positions 27 and 34 (K27M and G34R/V) in about 80% of diffuse intrinsic pontine glioma, a sub-group of pediatric high-grade glioma (pHGG).[61] Mechanistically, it has been reported that the K27M mutant heterotypic nucleosomes prevents PRC2 binding to the chromatin rather than sequestering the complex.[62,63] Notably, glioma cells carrying K27M have reduced H3K27^{me3} levels genome-wide leading to an increased expression of developmental genes normally repressed by PRC2.[62,63]

Recently, Jin et al.[40] proposed dual inhibition of BMI1 and EZH2 as a more efficient therapeutics against glioma than BMI1 (using PTC596) or EZH2 (using EPZ6438) inhibition alone. Drug regimen for PTC596 was once per week at 10 mg/kg. When PTC596 was used in combination with EPZ6438, they observed an increase in median lifespan of 13 days (PN19191 cell line) and of 17 days (Mes20 cell line) when compared to untreated animals, and of 12 and 3 days respectively when compared to PTC596-treated animals.[40] Considering that the plasma half-life of PTC596 is ~9 h at 10 mg/kg (with an AUC_{last} of ~30 h in the brain), it is predicted that the drug regimen used in these experiments resulted in lack of BMI1 inhibition for several days in between each treatment, allowing BMI1 reactivation in CSCs.

Herein, we found that when used at optimal concentration with a drug regimen allowing relatively constant BMI1 inhibition, PTC596 was able to extend median lifespan by 27 days (GBM1205 cell line) and 41 days (GBM0811 cell line) in terminally ill mice bearing tumors having either a mixed proneural/classical or a proneural phenotype. We found that PTC596 also rapidly impacted on EZH2, SOX2, and FOXG1 protein levels in vitro, possibly resulting in the modified tumor phenotype observed in some animals with relapse. In contrast, the A1016 inhibitor was generally more specific, showing modest effects on EZH2. However, because of its more rapid elimination from the circulation, A1016 was given daily. This resulted in lethal toxicity few days after the end of the treatment when A1016 was used at the highest concentration. Yet, three out four animals were tumor-free at necropsy. We conclude that A1016 would be the best candidate into the clinic for GBM treatment, but that more work is required to control peripheral toxicity in order to achieve high and stable brain concentration during the treatment period.

Interestingly, *BMI1* is overexpressed in cultured neurospheres from childhood brain tumors.[4] *BMI1* overexpression was also reported in 53% of pHGG in situ and *BMI1* inactivation in pHGG neurospheres impaired tumor formation in mouse xenografts.[64,65]

These findings thus indicate possible additional application for BMI1 inhibitors. A phase 1b study was initiated in 2018 for pHGG using PTC596 in combination with radiation therapy (NCT03605550). Considering that PTC596 can target both BMI1 and EZH2, it is recommended that in vitro and in vivo pre-clinical studies with pHGG cell lines be undertaken first to test for possible de-differentiation toward a more aggressive tumor phenotype. Our findings also raise concern that long-term treatment of solid tumors located outside the central nervous system with PTC596 could drive EMT, thus possibly resulting in a more aggressive and metastatic tumor phenotype.

METHODS

The complete method is available in Supplementary Information. The datasets generated during and/or analyzed during the current study are available from the corresponding author. All blots derived from the same experiment were processed in parallel. Primary GBM were obtained from the Department of Pathology of the Maisonneuve-Rosemont Hospital, and the Brain Tumor Tissue Bank (Toronto, ON, Canada). Fresh tumors were obtained from consenting patients and used with permission of our institutes' ethical committee (CER; Project No. 2007-402, 06099). The Maisonneuve-Rosemont Hospital's animal protection committee (CPA) approved all experiments performed in mice (Project No. 2015-23). Fresh GBM samples were processed for cell cultures within 1 h after reception. Tumor was washed and cut in small pieces before mechanical dissociation in oxygenated HBSS. Cell suspensions were passed over a 40-µm filter mesh. After centrifugation, cells were resuspended in GBM media: DMEM/F12 medium (Invitrogen) containing 0.25% glucose, N2 and B27 supplements, Heparin (2 µg/ml; Sigma), gentamicin (25 µg/ml; Invitrogen), human recombinant FGF2 (10 ng/ml; Peprotech), and human recombinant EGF (20 ng/ml; Sigma). Afterward, cultures were allowed to grow for 3 weeks to form spherical colonies (i.e., neurospheres). For passages, neurospheres were incubated in an enzyme-free solution (Millipore Bioscience Research Reagents) at 37 °C for 5 min, and mechanically dissociated with a 20G needle. After trituration, the cell suspension was plated in GBM media. Cell lines used in this study are: GBM1205, GBM1909, GBM0811, and GBM0410. Molecular weights for western blot analysis are showed with the original blots in Supplementary Information as Supplementary Figs. 10 and 11. BMI1 inactivation using CRISPR/Cas9 were carried out by the polymeric delivery of a Cas9-expressing plasmid (Dharmacon #CAS10140), a synthetic guide RNA (sgRNA) Scramble (Dharmacon #U-007501) or complementary to BMI1 (Target: AACGTGTATTGTTCGTTACC) and a synthetic trans-activating crRNA (Dharmacon #U-002005) using Mirus TransIT-X2 (Cat#MIR6003) according to the manufacturer's instructions. To enhance the knockout efficiency, a full six-well plate of neurospheres was dissociated and condensed in 1 ml of GBM media and platted onto a well of an ultra-low attachment six-well plate.

Reporting Summary

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

DATA AVAILABILITY

Raw data, cell lines, and reagents are available upon request.

CODE AVAILABILITY

Accession code for RNA-seq data presented in Figs. 4 and 5 is GSE136829.

Received: 24 February 2019; Accepted: 25 October 2019;
Published online: 06 January 2020

REFERENCES

1. Holland, E. C. Gliomagenesis: genetic alterations and mouse models. *Nat. Rev. Genet.* **2**, 120–129 (2001).
2. Louis, D. N. et al. The 2007 WHO classification of tumours of the central nervous system. *Acta Neuropathol.* **114**, 97–109 (2007).
3. Maher, E. A. et al. Malignant glioma: genetics and biology of a grave matter. *Genes Dev.* **15**, 1311–1333 (2001).
4. Hemmati, H. D. et al. Cancerous stem cells can arise from pediatric brain tumors. *Proc. Natl Acad. Sci. USA* **100**, 15178–15183 (2003).
5. Singh, S. K. et al. Identification of a cancer stem cell in human brain tumors. *Cancer Res.* **63**, 5821–5828 (2003).
6. Singh, S. K. et al. Identification of human brain tumour initiating cells. *Nature* **432**, 396–401 (2004).
7. Lee, J. et al. Tumor stem cells derived from glioblastomas cultured in bFGF and EGF more closely mirror the phenotype and genotype of primary tumors than do serum-cultured cell lines. *Cancer Cell* **9**, 391–403 (2006).
8. Bao, S. et al. Glioma stem cells promote radioresistance by preferential activation of the DNA damage response. *Nature* **444**, 756–760 (2006).
9. Facchino, S. et al. Brain cancer stem cells: current status on glioblastoma multiforme. *Cancers (Basel)* **3**, 1777–1797 (2011).
10. Facchino, S. et al. BMI1 confers radioresistance to normal and cancerous neural stem cells through recruitment of the DNA damage response machinery. *J. Neurosci.* **30**, 10096–10111 (2010).
11. Ligon, K. L. et al. Olig2-regulated lineage-restricted pathway controls replication competence in neural stem cells and malignant glioma. *Neuron* **53**, 503–517 (2007).
12. Wang, Y. et al. Expression of mutant p53 proteins implicates a lineage relationship between neural stem cells and malignant astrocytic glioma in a murine model. *Cancer Cell* **15**, 514–526 (2009).
13. Chudnovsky, Y. et al. ZFHX4 interacts with the NuRD core member CHD4 and regulates the glioblastoma tumor-initiating cell state. *Cell Rep.* **6**, 313–324 (2014).
14. Liu, H. K. et al. The nuclear receptor tailless induces long-term neural stem cell expansion and brain tumor initiation. *Genes Dev.* **24**, 683–695 (2010).
15. Gallo, M. et al. MLL5 orchestrates a cancer self-renewal state by repressing the histone variant H3.3 and globally reorganizing chromatin. *Cancer Cell* **28**, 715–729 (2015).
16. Suva, M. L. et al. Reconstructing and reprogramming the tumor-propagating potential of glioblastoma stem-like cells. *Cell* **157**, 580–594 (2014).
17. Sparmann, A. & van Lohuizen, M. Polycomb silencers control cell fate, development and cancer. *Nat. Rev. Cancer* **6**, 846–856 (2006).
18. Levine, S. S. et al. The core of the polycomb repressive complex is compositionally and functionally conserved in flies and humans. *Mol. Cell. Biol.* **22**, 6070–6078 (2002).
19. Dellino, G. I. et al. Polycomb silencing blocks transcription initiation. *Mol. Cell* **13**, 887–893 (2004).
20. Kuzmichev, A. et al. Histone methyltransferase activity associated with a human multiprotein complex containing the Enhancer of Zeste protein. *Genes Dev.* **16**, 2893–2905 (2002).
21. Wang, H. et al. Role of histone H2A ubiquitination in Polycomb silencing. *Nature* **431**, 873–878 (2004).
22. Kalb, R. et al. Histone H2A monoubiquitination promotes histone H3 methylation in Polycomb repression. *Nat. Struct. Mol. Biol.* **21**, 569–571 (2014).
23. Kleer, C. G. et al. EZH2 is a marker of aggressive breast cancer and promotes neoplastic transformation of breast epithelial cells. *Proc. Natl Acad. Sci. USA* **100**, 11606–11611 (2003).
24. Lassman, A. B. et al. Overexpression of c-MYC promotes an undifferentiated phenotype in cultured astrocytes and allows elevated Ras and Akt signaling to induce gliomas from GFAP-expressing cells in mice. *Neuron Glia Biol.* **1**, 157–163 (2004).
25. Orian, J. M. et al. Overexpression of multiple oncogenes related to histological grade of astrocytic glioma. *Br. J. Cancer* **66**, 106–112 (1992).
26. Valk-Lingbeek, M. E. et al. Stem cells and cancer; the polycomb connection. *Cell* **118**, 409–418 (2004).
27. Varambally, S. et al. The polycomb group protein EZH2 is involved in progression of prostate cancer. *Nature* **419**, 624–629 (2002).
28. 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).
29. Abdouh, M. et al. BMI1 sustains human glioblastoma multiforme stem cell renewal. *J. Neurosci.* **29**, 8884–8896 (2009).
30. Dimri, G. P. et al. The Bmi-1 oncogene induces telomerase activity and immortalizes human mammary epithelial cells. *Cancer Res.* **62**, 4736–4745 (2002).
31. Haupt, Y. et al. Novel zinc finger gene implicated as myc collaborator by retrovirally accelerated lymphomagenesis in E mu-myc transgenic mice. *Cell* **65**, 753–763 (1991).
32. Jacobs, J. J. et al. Bmi-1 collaborates with c-Myc in tumorigenesis by inhibiting c-Myc-induced apoptosis via INK4a/ARF. *Genes Dev.* **13**, 2678–2690 (1999).

33. Leung, C. et al. Bmi1 is essential for cerebellar development and is overexpressed in human medulloblastomas. *Nature* **428**, 337–341 (2004).
34. van Lohuizen, M. et al. Sequence similarity between the mammalian bmi-1 proto-oncogene and the Drosophila regulatory genes Psc and Su(z)2. *Nature* **353**, 353–355 (1991).
35. van Lohuizen, M. et al. Identification of cooperating oncogenes in E mu-myc transgenic mice by provirus tagging. *Cell* **65**, 737–752 (1991).
36. Bruggeman, S. W. et al. Bmi1 controls tumor development in an Ink4a/Arf-independent manner in a mouse model for glioma. *Cancer Cell* **12**, 328–341 (2007).
37. Moon, J. H. et al. Induction of neural stem cell-like cells (NSCLCs) from mouse astrocytes by Bmi1. *Biochem. Biophys. Res. Commun.* **371**, 267–272 (2008).
38. Chatoo, W. et al. Bmi1 distinguishes immature retinal progenitor/stem cells from the main progenitor cell population and is required for normal retinal development. *Stem Cells* **28**, 1412–1423 (2010).
39. Verhaak, R. G. et al. Integrated genomic analysis identifies clinically relevant subtypes of glioblastoma characterized by abnormalities in PDGFRA, IDH1, EGFR, and NF1. *Cancer Cell* **17**, 98–110 (2010).
40. Jin, X. et al. Targeting glioma stem cells through combined BMI1 and EZH2 inhibition. *Nat. Med.* **23**, 1352–1361 (2017).
41. Nishida, Y. et al. The novel BMI-1 inhibitor PTC596 downregulates MCL-1 and induces p53-independent mitochondrial apoptosis in acute myeloid leukemia progenitor cells. *Blood Cancer J.* **7**, e527 (2017).
42. Seoane, J. et al. Integration of Smad and forkhead pathways in the control of neuroepithelial and glioblastoma cell proliferation. *Cell* **117**, 211–223 (2004).
43. Sturm, D. et al. Hotspot mutations in H3F3A and IDH1 define distinct epigenetic and biological subgroups of glioblastoma. *Cancer Cell* **22**, 425–437 (2012).
44. Bulstrode, H. et al. Elevated FOXG1 and SOX2 in glioblastoma enforces neural stem cell identity through transcriptional control of cell cycle and epigenetic regulators. *Genes Dev.* **31**, 757–773 (2017).
45. Gangemi, R. M. et al. SOX2 silencing in glioblastoma tumor-initiating cells causes stop of proliferation and loss of tumorigenicity. *Stem Cells* **27**, 40–48 (2009).
46. Suva, M. L. et al. EZH2 is essential for glioblastoma cancer stem cell maintenance. *Cancer Res.* **69**, 9211–9218 (2009).
47. Chatoo, W. et al. The polycomb group gene Bmi1 regulates antioxidant defenses in neurons by repressing p53 pro-oxidant activity. *J. Neurosci.* **29**, 529–542 (2009).
48. Flamier, A. et al. Modeling late-onset sporadic Alzheimer's disease through BMI1 deficiency. *Cell Rep.* **23**, 2653–2666 (2018).
49. 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).
50. Abdouh, M. et al. The Polycomb Repressive Complex 1 protein BMI1 is required for constitutive heterochromatin formation and silencing in mammalian somatic cells. *J. Biol. Chem.* **291**, 182–197 (2015).
51. Pinheiro, I. et al. Prdm3 and Prdm16 are H3K9me1 methyltransferases required for mammalian heterochromatin integrity. *Cell* **150**, 948–960 (2012).
52. Boyer, L. A. et al. Polycomb complexes repress developmental regulators in murine embryonic stem cells. *Nature* **441**, 349–353 (2006).
53. Cao, R. et al. Role of Bmi-1 and Ring1A in H2A ubiquitylation and Hox gene silencing. *Mol. Cell* **20**, 845–854 (2005).
54. Coskun, V. et al. CD133+ neural stem cells in the ependyma of mammalian postnatal forebrain. *Proc. Natl Acad. Sci. USA* **105**, 1026–1031 (2008).
55. Lee, A. et al. Isolation of neural stem cells from the postnatal cerebellum. *Nat. Neurosci.* **8**, 723–729 (2005).
56. Takahashi, K. et al. Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell* **131**, 861–872 (2007).
57. Singh, D. K. et al. Oncogenes activate an autonomous transcriptional regulatory circuit that drives glioblastoma. *Cell Rep.* **18**, 961–976 (2017).
58. Kim, S. H. et al. EZH2 protects glioma stem cells from radiation-induced cell death in a MELK/FOXM1-dependent manner. *Stem Cell Rep.* **4**, 226–238 (2015).
59. Smits, M. et al. miR-101 is down-regulated in glioblastoma resulting in EZH2-induced proliferation, migration, and angiogenesis. *Oncotarget* **1**, 710–720 (2010).
60. de Vries, N. A. et al. Prolonged Ezh2 depletion in glioblastoma causes a robust switch in cell fate resulting in tumor progression. *Cell Rep.* **10**, 383–397 (2015).
61. Wu, G. et al. Somatic histone H3 alterations in pediatric diffuse intrinsic pontine gliomas and non-brainstem glioblastomas. *Nat. Genet.* **44**, 251–253 (2012).
62. Chan, K. M. et al. The histone H3.3K27M mutation in pediatric glioma reprograms H3K27 methylation and gene expression. *Genes Dev.* **27**, 985–990 (2013).
63. 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).
64. Baxter, P. A. et al. Silencing BMI1 eliminates tumor formation of pediatric glioma CD133+ cells not by affecting known targets but by down-regulating a novel set of core genes. *Acta Neuropathol. Commun.* **2**, 160 (2015).
65. Kumar, S. S. et al. BMI-1 is a potential therapeutic target in diffuse intrinsic pontine glioma. *Oncotarget* **8**, 62962–62975 (2017).

ACKNOWLEDGEMENTS

This work was supported by grants from the Cancer Research Society (CRS), National Science and Engineering Research Council of Canada (NSERC) and PTC Therapeutics. A.F., R.H., and A.B. were supported by fellowships from the Molecular Biology Program of Université de Montréal. We thank PTC Therapeutics for the generous gift of the PTC596 and A1016 compounds and for providing critical information on the drug's pharmacokinetics and serum concentration in orally treated mice.

AUTHOR CONTRIBUTIONS

Conceived and designed: G.B. Performed the experiments: A.F., M.A., R.H., A.B. and A.G. Analyzed the data: G.B., A.F., R.H., A.B. and M.A. Wrote the paper: G.B. and A.F.

COMPETING INTERESTS

The author(s) declare no competing interests. PTC Therapeutics was not involved in the design of the study, analysis and interpretation of the data, or writing of the manuscript.

ADDITIONAL INFORMATION

Supplementary information is available for this paper at https://doi.org/10.1038/s41698-019-0106-1.

Correspondence and requests for materials should be addressed to G.B.

Reprints and permission information is available at http://www.nature.com/reprints

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

10 Annex V: Deregulation of Neuro-Developmental Genes and Primary Cilium Cytoskeleton Anomalies in iPSC Retinal Sheets from Human Syndromic Ciliopathies

This fourth annex is an article, on which I have worked during my Ph.D., and was published on Stem Cell Reports in 2020. This article is not in the main aim of my studies, nevertheless, it is an important article establishing a new model for the study of retinal degeneration diseases.

Stem Cell Reports
Article

ISSCR

OPEN ACCESS

Deregulation of Neuro-Developmental Genes and Primary Cilium Cytoskeleton Anomalies in iPSC Retinal Sheets from Human Syndromic Ciliopathies

Andrea Barabino,[1] Anthony Flamier,[1,6] Roy Hanna,[1] Elise Héon,[2] Benjamin S. Freedman,[3,*] and Gilbert Bernier[1,4,5,7,*]

[1]Stem Cell and Developmental Biology Laboratory, Hôpital Maisonneuve-Rosemont, 5415 Boul. Assomption, Montreal, QC H1T 2M4, Canada
[2]Hospital for Sick Children, Department of Ophthalmology and Vision Sciences, Program of Genetics and Genome Biology, 555 University av., Toronto, ON M5G 1X8, Canada
[3]Department of Medicine, Division of Nephrology, Kidney Research Institute, and Institute of Stem Cell and Regenerative Medicine, and Department of Pathology, University of Washington School of Medicine, Seattle, WA 98109, USA
[4]Department of Neurosciences, University of Montreal, Montreal, QC H3C 3J7, Canada
[5]Department of Ophthalmology, University of Montreal, Montreal, QC H3C 3J7, Canada
[6]Present address: Whitehead Institute for Biomedical Research, Massachusetts Institute of Technology (MIT), 455 Main Street, Cambridge, MA 02142, USA
[7]Lead Contact
*Correspondence: benof@uw.edu (B.S.F.), gbernier.hmr@ssss.gouv.qc.ca (G.B.)
https://doi.org/10.1016/j.stemcr.2020.02.005

SUMMARY

Ciliopathies are heterogeneous genetic diseases affecting primary cilium structure and function. Meckel-Gruber (MKS) and Bardet-Biedl (BBS) syndromes are severe ciliopathies characterized by skeletal and neurodevelopment anomalies, including polydactyly, cognitive impairment, and retinal degeneration. We describe the generation and molecular characterization of human induced pluripotent stem cell (iPSC)-derived retinal sheets (RSs) from controls, and MKS (TMEM67) and BBS (BBS10) cases. MKS and BBS RSs displayed significant common alterations in the expression of hundreds of developmental genes and members of the WNT and BMP pathways. Induction of crystallin molecular chaperones was prominent in MKS and BBS RSs suggesting a stress response to misfolded proteins. Unique to MKS photoreceptors was the presence of supernumerary centrioles and cilia, and aggregation of ciliary proteins. Unique to BBS photoreceptors was the accumulation of DNA damage and activation of the mitotic spindle checkpoint. This study reveals how combining cell reprogramming, organogenesis, and next-generation sequencing enables the elucidation of mechanisms involved in human ciliopathies.

INTRODUCTION

Primary cilia are typically non-motile cytoplasmic extensions of a microtubule-based structure that projects from the cell surface and are indispensable for normal developmental and physiological functions (Nonaka et al., 1998; Satir et al., 2007). The ciliary axoneme develops from and is anchored to a specialized centriole called the basal body (BB) that acts as a microtubule organizing center. The BB is a symmetric radial arrangement of nine triplet microtubules from which the outer doublet of microtubules of the axoneme extends. Syndromic and non-syndromic ciliopathies represent a group of heterogeneous genetic diseases caused by mutations affecting the structure and function of the primary cilium. Phenotypic and genetic heterogeneity is frequently observed in these diseases. The relationship between cilia genes and ciliopathies is far more complex than that described by classical Mendelian genetics and has recently been the focus of numerous reviews (Loncarek and Bettencourt-Dias, 2018; Vertii et al., 2015; Wheway et al., 2018).

Meckel-Gruber syndrome (MKS) and Bardet-Biedl syndrome (BBS) are rare developmental diseases characterized by multiple developmental anomalies, including retinal degeneration, digits and genito-urinary defects, as well as cognitive impairment (Leitch et al., 2008; Álvarez-Satta et al., 2017). Although MKS is fatal at birth, BBS is one of the most severe ciliopathies that is compatible with life. MKS is a lethal autosomal recessive ciliopathy, also presenting polycystic kidneys and severe eye/brain malformations, with over 13 disease-associated genes (Delous et al., 2007; Kyttälä et al., 2006; Smith et al., 2006; OMIM, 2019a). In contrast, BBS is a viable disorder associated with obesity and retinal degeneration (Leitch et al., 2008; Álvarez-Satta et al., 2017) and with variants identified in over 22 genes, with *BBS10* representing the most commonly mutated gene (Nishimura et al., 2004; OMIM, 2019b). MKS and BBS were in principle considered as two distinct clinical entities. However, the identification of hypomorphic mutations in *MKS1* and *TMEM67* in some BBS patients introduced the concept that BBS may represent a milder form of MKS (Leitch et al., 2008). The MKS3 protein (encoded by *TMEM67*) is apparently not an integral part of the MKS complex, but more likely to interact with it (Leitch et al., 2008; Smith et al., 2006). MSK3 localizes to the transition zone at the base of the primary cilium, and

gulated genes

to the plasma membrane in ciliated cells (Reiter et al., 2012). MKS3 has been associated with centrosome migration to the apical cell surface during early ciliogenesis and with the regulation of centrosome organization (Adams et al., 2012; Gupta et al., 2015). The BB derives from the mother centriole, which distinguishes it from the daughter centriole by the presence of appendages on its distal end, allowing the formation of the primary cilium (Chang et al., 2003). Centriolin (also known as CEP110 or CNTRL) is associated with the mother centriole (Ou et al., 2002). Mutations in *CNTRL* can cause ciliopathies as atrioventricular septal defect (Burnicka-Turek et al., 2016). Although many BBS-associated proteins are an integral part of the BBSome complex, BBS10 most likely interacts with the BBSome and shows sequence similarities with chaperones (Stoetzel et al., 2006).

The vast majority of genes causing retinal degeneration involves a defect in a ciliary protein (RetNet). Molecular and cellular insights found in this work are potentially relevant to other ciliopathies and non-syndromic retinal disorders. Degeneration of photoreceptors (PRs) is often part of syndromic ciliopathies (Adams et al., 2007; Novarino et al., 2011). Retinal degeneration varies depending on the type of PRs that are primarily affected. Rods respond to dim light and are important for night and peripheral vision. Cones respond to intense light and are required for color, daylight, and high-resolution central vision (Aboshiha et al., 2016). The inner segment (IS) and outer segment (OS) of PRs, which respectively represent the metabolic factory and the structure where phototransduction takes place, are connected by the connecting cilium (CC), which is a modified intracellular version of the primary cilia (Young, 1968). The proteins synthesized in the IS are transported to the base of the CC in post-Golgi vesicles in an area also called the transition zone, where they are associated with molecular transport complexes that allow bidirectional movement along the axoneme, called intraflagellar transport (IFT) (Khanna, 2015; Pazour et al., 2002). The IFT family and other BB-associated proteins, if mutated, may cause mislocalization and accumulation of OS proteins in the IS, causing PR degeneration (Marszalek et al., 2000; Pazour et al., 2002). The RPGR, RP1, and RP2 proteins mostly localize at the BB of the CC, and mutations in these are a leading cause of retinitis pigmentosa (Breuer et al., 2002).

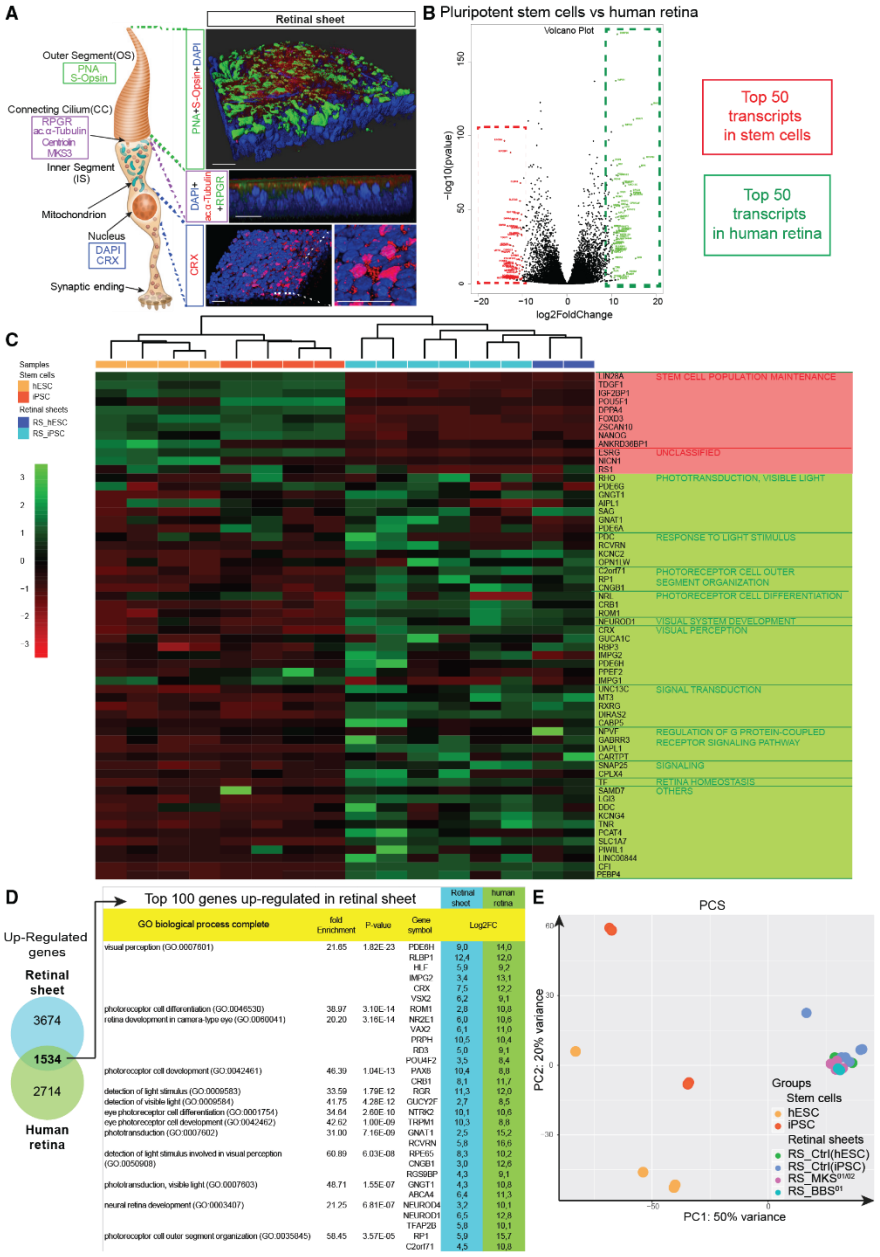
Although valuable, animal models have shown limitations in modeling retinal ciliopathies, with patient-derived induced pluripotent stem cells (iPSCs) representing a new opportunity for modeling human diseases (Grandy et al., 2019). Structures like the macula, responsible for high-resolution central vision, are not present in most commonly used animal models. Moreover, the PR's ultrastructure between mice and humans differs at the transition zone, an area adjacent to the CC. Here, we found structures called calyceal processes which are present in human but not in mice. Likewise, the USH1G/1C proteins, which are mutated in Usher syndrome, localize to the CC of human PRs but have no equivalent in rodents (Sahly et al., 2012). We report here on the generation of iPSCs from MKS and BBS cases and on their differentiation into retinal sheets (RSs) containing cone PRs (Zhou et al., 2015). We found that RSs from MKS and BBS cases displayed common alterations in hundreds of developmental genes, including homeobox and *HOX* genes, as well as genes of the WNT, NOTCH, and BMP signaling pathways. MKS PRs showed supernumerary cilia and centrioles and mislocalization of ciliary proteins. BBS PRs presented mitotic spindle checkpoint activation, DNA damage, and genomic instability. Both MKS and BBS PRs accumulated chaperones of the crystallin gene family, suggesting a cellular response to misfolded proteins and/or proteasome dysfunction. This study brings new molecular and cell biological information on the neurodevelopmental and retinal degeneration anomalies associated with human syndromic ciliopathies.

RESULTS

Generation and Characterization of iPSC-Derived RSs

We have generated iPSCs from skin fibroblasts of three healthy volunteers ($Ctrl^{01}$, $Ctrl^{02}$, and $Ctrl^{03}$), two unrelated MKS cases (MKS^{01} and MKS^{02}), and two unrelated BBS cases (BBS^{01} and BBS^{02}). The iPSCs expressed pluripotency markers, were able to form teratomas and could be differentiated into embryoid bodies containing retinal pigment epithelium (RPE) upon exposure to nicotinamide (Figures S1A–S1D). The presence of an external cilium was also observed in iPSCs using scanning electron microscopy (Figure S1B). By whole-genome sequencing (WGS) of the iPSC lines, we found that both BBS cases carried different mutations in *BBS10*, with a homozygous mutation causing a frameshift stop in the first case (c.271dup), and a compound heterozygous mutation in the second case (c.909_912del; c.687del) (Figures S1E–S1G). The MKS^{01} case carried compound heterozygous permutations at the *TMEM67* locus (c.233G > A; c.1046T > C) predicted to be pathogenic (Figures S1E–S1G). We could not, however, confirm the disease-causing mutation in the MKS^{02} case.

Control, $BBS^{01/02}$, and $MKS^{01/02}$ iPSCs were differentiated for 60 days *in vitro* (DIV) into RSs using the same methodology as for human embryonic stem cells (hESCs) (Zhou et al., 2015). Control RSs were analyzed by confocal immunofluorescence. 3D reconstruction imaging revealed the generation of a polarized, multi-layered tissue expressing OS (peanut agglutinin [PNA] and S-opsin), CC (acetylated



(legend on next page)

α-tubulin and RPGR), and nuclear pan-PR (CRX) markers (Figure 1A), suggesting efficient differentiation into RSs. Using RNA sequencing (RNA-seq), we compared the transcriptome of control RSs with that of the Human Retinal Development Atlas (Hoshino et al., 2017), and found that RSs clustered with human embryonic retinas at days 80 and 94. RNA-seq data from control RSs were also compared with the top 50 most upregulated transcripts in adult human retina and the top 50 iPSC-specific transcripts (Li et al., 2014) (Figure 1B). Retinal and PR-specific transcripts, including *NR2E1*, *GRK1*, *CRX*, *PDE6H*, *RXRG*, *ROM1*, *CRB1*, *RPE65*, *PRPH*, *RP1*, *RCVRN*, *ABCA4*, and *GNAT1*, were significantly enriched in RSs. In contrast, iPSC-specific transcripts were lost in RSs, suggesting efficient cellular differentiation (Figure 1C). Despite the presence of rod-specific transcripts in RSs, we failed to detect rhodopsin protein expression using immunoblot or immunofluorescence.

When comparing the entire gene dataset of undifferentiated iPSCs with that of adult human retinas and RSs, we found that 1,534 genes were commonly upregulated in human retinas and RSs. Gene ontology analysis of the top 100 genes revealed that these were primarily implicated in the detection of light stimuli, PR OS organization, PR development and visual perception (Figure 1D). We also performed a principal-component analysis of RSs and iPSC/hESC lines. Although iPSCs and hESCs were generally scattered, RSs from all samples were grouped together, suggesting cellular differentiation toward a common lineage (Figure 1E). We next used immunoblot to test if RSs expressed PR-specific proteins. We found that $Ctrl^{01/02/03}$, $MKS^{01/02}$, and $BBS^{01/02}$ RSs were all expressing CRX (expressed in PR and in a subset of bipolar neurons) and S-opsin (expressed exclusively in cone PRs) (Figure 2A). Taken as a whole, these results suggested robust differentiation of iPSC lines into RSs.

Deregulation of Neurodevelopmental Genes in MKS and BBS RSs

We compared RNA-seq data from $Ctrl^{01/02}$, $MKS^{01/02}$, and BBS^{01} RSs and found thousands of genes commonly dysregulated in ciliopathies (Figures 2B and 2C). A total of 1,597 genes were dysregulated in both BBS^{01} and $MKS^{01/02}$ samples when compared with controls (661 upregulated, 936 downregulated) (Figures 2D and 3A). Among the 661 commonly upregulated genes, most were related to neural/retinal development (*DLX1*, *VSX1*, and *SIX6*), differentiation (*NEUROD4* and *ASCL1/MASH1*), and function (*SLC32A1* and *NTRK1*), while others were associated with ganglion and amacrine cell fate, suggesting premature and/or increased neurogenesis and perturbed retinal cell fate specification in ciliopathies (Figures 2E–2G and S2A–S2D). Different homeobox genes including members of the HOX family were also upregulated in both BBS^{01} and $MKS^{01/02}$ RSs (Figure S2A). Defective NOTCH and/or WNT signaling is frequently associated with premature neurogenesis and/or cell-cycle exit of neural progenitors (Ma et al., 2019; Navarro Quiroz et al., 2018). Hence, we found altered expression of genes of the WNT and NOTCH signaling pathways in both MKS and BBS RSs (Figures S2B and S2C). Further analyses suggested major deregulation of WNT signaling in ciliopathies toward the canonical pathway at the expense of the non-canonical one, which is known to play a role in planar cell polarity and cilia formation (Figure S2B) (May-Simera and Kelley, 2012).

Alterations of Genes Involved in Development, Morphogenesis, and Cilia Formation

When analyzing genes commonly downregulated in BBS^{01} and $MKS^{01/02}$ RSs, we found significant alteration of genes involved in development (112 genes) and/or morphogenesis (138 genes) (Figures 3A–3C). We next grouped these

Figure 1. Generation of iPSC-Derived Retinal Sheets

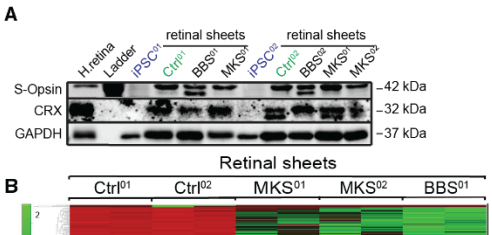
(A) Representation of a cone PRs (left) and 3D reconstruction of an RS (right) with reference to the main compartments, outer segment (OS), connecting cilium (CC), inner segment (IS), and nucleus. For each compartment are indicated the main markers used in this paper: peanut agglutinin (PNA) and S-opsin for the OS; acetyl-α-tubulin and RPGR for the CC; CRX for the nucleus. Scale bars, 40 μm.

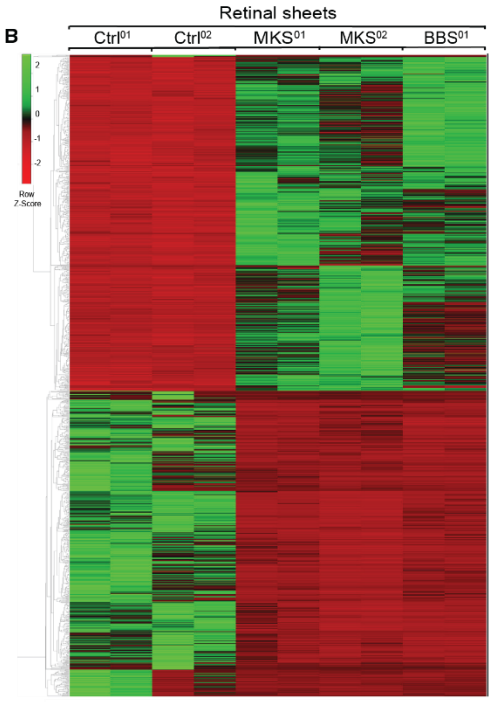
(B) Volcano plot from RNA-seq analyses between human retina (n = 2 independent biological samples) and undifferentiated stem cells (n = 4 iPSCs; n = 4 hESCs, independent cell lines). The red points show the 50 most significantly downregulated genes (stem cell-related genes). The green points show the 50 most significantly upregulated genes (human retina-related genes).

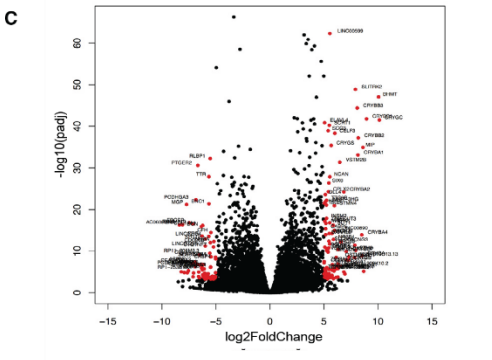
(C) Heatmap showing differential expression for stem cell and human retina-related genes between undifferentiated stem cells lines (n = 4 iPSCs; n = 4 hESCs) and Ctrl RSs (n = 8 from 3 independent iPSC and 1 hESC lines). Hierarchical clustering by Pearson's correlation clearly distinguishes the two groups.

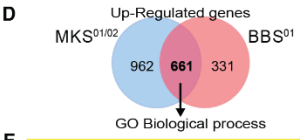
(D) Venn diagrams showing the intersection of significant genes differentially upregulated in control RSs ($Ctrl^{01}$, $Ctrl^{02}$, $Ctrl^{03}$, and hESC-Ctrl) and in human retina. Gene ontology enrichment analysis of the 100 most upregulated genes in RSs compared with undifferentiated iPSCs that are also upregulated in the human retina ($p < 0.05$). For every gene ontology (GO) group, we indicated their fold enrichment, p values, and few representative genes. For every gene are indicated log2FC in Ctrl_RSs (light blue) and in human retina (green).

(E) Principal-component analysis (PCA) between stem cells (iPSCs and ESCs) and RSs from control iPSCs and hESCs, and RS from MKS (n = 4 from 2 patients) and BBS (n = 2 from 1 patient). Notice how all RSs cluster together, while PSCs are more dispersed.









E

Neural retina function, development and morphogenesis	Gene Symbol	Log2FC MKS	Log2FC BBS
visual perception(GO:0007601)	GJA8	6,5	4,7
	GJA3	2,1	2,1
	RGS16	3,0	2,1
	IMPG2	4,6	3,1
	GUCY2D	3,6	3,1
	CHRNB2	4,1	3,5
	ZIC2	2,2	2,4
	MIP	8,4	8,7
	COL2A1	2,3	2,5
	VSX2	4,2	5,2
	RBP3	4,9	4,4
	CACNA1F	3,2	2,8
	PDC	4,6	4,4
	CLRN1	2,7	3,2
	RAX2	5,3	3,9
	NDP	5,3	3,2
	VSX1	5,5	4,6
	CNGA3	4,7	3,8
	RCVRN	4,3	4,3
	KCNJ10	3,8	4,4
phototransduction, visible light (GO:0007603)	AIPL1	3,9	3,3
	PDE6C	3,4	2,3
retinal cone cell development (GO:0046549	RORB	3,5	3,9
	DIO3	3,9	2,2
eye photoreceptor cell development (GO:0042462)	CRB1	3,3	3,6
	PRDM1	4,1	4,5
	TH	3,1	2,8
retina development in camera-type eye (GO:0060041)	GRM6	4,6	3,7
	NR2E1	6,3	4,6
	POU4F2	4,5	5,8
regulation of neural retina development (GO:0061074)	DLL4	4,6	5,8
	PTF1A	4,5	5,7
neural retina development (GO:0003407)	ATOH7	4,6	6,4
	SLC17A7	3,4	3,3
eye development (GO:0001654)	FOXE3	7,2	6,8
	SIX6	5,3	5,8
	SOX2	3,1	2,2
	RBP4	3,5	3,7
	CHRDL1	3,6	2,3
camera-type eye development (GO:0043010)	HES5	4,6	5,6
	PAX2	8,4	3,6
camera-type eye morphogenesis (GO:0048593)	AQP5	2,0	2,2

F

Retinal ganglion cell development	Gene Symbol	Log2FC MKS	Log2FC BBS
retinal ganglion cell axon guidance (GO:0031290)	ISL2	5,3	5,3
	ISL1	5,1	4,6
	SLIT1	6,3	3,4
	POU4F2	4,5	5,8
	NRCAM	3,9	2,7
optic nerve development (GO:0021554)	ATOH7	4,6	6,4
	PAX2	8,4	3,6
	NTRK1	3,4	4,2
	DRGX	3,1	3,6
	POU4F1	4,4	3,2

G

Amacrine cell development	Gene Symbol	Log2FC MKS	Log2FC BBS
amacrine cell differentiation (GO:0035881)	NEUROD4	5,4	4,9
	NEUROD1	4,8	5,3
	FOXN4	3,6	4,4
	PTF1A	4,5	5,7
positive regulation of amacrine cell differentiation (GO:1902871)	DLX1	6,7	4,6
	DLX2	4,9	3,4

(legend on next page)

genes among the most representative clusters and compared them by focusing on genes common to different organs/systems (Figures 3A–3D). We found five developmental genes—*Noggin* (*NOG*), *SIX1*, *BMP4*, *CDH19*, *OSR1*—all involved in kidney, heart, ear, skeletal, circulatory, and retinal development (Ahmed et al., 2012; Wu et al., 2014). Three of them, *NOG*, *SIX1*, and *OSR1*, were also present in most of the morphogenetic groups. Other genes, such as *MSX1*, *TWIST1*, *SIX2*, and *OSR2* were present in some of the morphogenetic and developmental groups (Figure 3B) (Goodnough et al., 2016; Li et al., 2003; Paradowska-Stolarz, 2015). *PITX2* was one of the most downregulated developmental genes in both $MKS^{01/02}$ and BBS^{01} RSs (Figure 3D). Notably, a large number of genes involved in cilium assembly/organization and intraciliary transport were downregulated in MKS and BBS RSs (Figures S3A–S3C) (HUGO, 2019). From these, 21 corresponded to the cilia- and flagella-associated protein gene family (Figure S3B). We also found 962 genes upregulated and 684 genes downregulated only in $MKS^{01/02}$ RSs (Figures S3D–S3H). Among these, many were associated with the development and anterior-posterior pattern specification (Figures S3E and S3H). Several genes of the HOX family were also specifically upregulated only in $MKS^{01/02}$ RSs, in agreement with MKS being the most severe form of syndromic ciliopathy (Figure S3F).

Induction of the Crystallin Molecular Chaperones in MKS and BBS RSs

From the top 10 most upregulated genes common to BBS and MKS RSs, we found that six encoded members of the crystallin gene family (Figures S4A–S4C) (Kamachi et al., 2001). In non-lens tissues, these proteins work as molecular chaperones against protein misfolding. Using immunoblot and immunofluorescence, we observed specific accumulation of CRYBB1, CRYBB2, and CRYBB3 in both the nuclear and cytoplasmic cell compartments of BBS^{01} and MKS^{01} PRs at DIV45 (Figures 4A–4C, S4A, and S4B). Polyubiquitin chains at lysine 48 (polyUb-K48) mark proteins for proteasomal degradation and poly-ubiquitylated proteins frequently accumulate in neurodegenerative diseases (Smith et al., 2015). When compared with control RSs, MKS^{01} and MKS^{02}, but not BBS^{01} RSs, presented increased polyUb-K48 levels, suggesting perturbed proteostasis (Figure 4B). Accumulation of misfolded proteins is frequently associated with neuronal cell death (Tzekov et al., 2011). Accordingly, we observed increased expression of apoptosis and cell death-related genes in $MKS^{01/02}$ and BBS^{01} RSs (Figure S4D). Induction of PR cell death in MKS and BBS RSs was also confirmed using the TUNEL assay and transmission electron microscopy (TEM) (Figures 4D–4I and S4E–S4H). These results suggested a cellular response to misfolded proteins and increased cell death in $MKS^{01/02}$ and BBS^{01} PRs.

MKS PRs Are Characterized by the Presence of Supernumerary Centrioles

We performed cell biological analyses to study the structure of the cilium in MKS and BBS PRs. The centriole forms the basement of the primary cilium, and Centriolin marks the mother centriole and BB (Chang et al., 2003). Using an antibody against Centriolin, we found that, when compared with controls, the number of centrioles per cell was higher in DIV60 $MKS^{01/02}$ PRs, but not in $BBS^{01/02}$ PRs (Figures 5A and 5B). This was confirmed by TEM (Figures 5C and 5D), where cilia also appeared to be shorter in MKS (Figures 5C and 5D). Although the total number of cells tended to be higher in MKS and BBS cultures, the difference was not significant (not shown). Notably, supernumerary centrioles were not observed in undifferentiated MKS^{01} iPSCs (Figures S5A–S5C), suggesting that the defect was cell-type specific.

We did not elucidate the pathogenic mutation in the MKS^{02} case. However, in the MKS^{01} case, the two identified mutations in *TMEM67* were predicted to generate a pathogenic but full-length variant of MKS3 (Figures S1G and

Figure 2. Neurodevelopmental Anomalies in MKS and BBS RSs

(A) Immunoblot on extracts from a human retina (positive ctrl), undifferentiated iPSCs (negative ctrl), and RSs from $Ctrl^{01/02}$, $BBS^{01/02}$, and $MKS^{01/02}$ patients.

(B) Heatmap of gene expression profile for differentially expressed genes between $Ctrl^{01/02}$ (n = 4 from 2 independent ctrl cell lines), $MKS^{01/02}$ (n = 4 from two independent MKS patients), and BBS^{01} (n = 2 from one BBS patient) ($p < 0.0001$: up, log2FC > 2; down, log2FC < −2). Hierarchical clustering by average linkage with Kendall's Tau correlation clearly distinguishes control and patient RSs.

(C) Volcano plot from RNA-seq analyses between $Ctrl^{01/02}$ and all patient RSs. The red points show the most significantly dysregulated genes (log2FC > 5 and adjusted $p < 0.001$).

(D) Venn diagrams showing the intersection of significant genes differentially upregulated in $MKS^{01/02}$ (blue, 1,623 genes) and in BBS^{01} RSs (red, 992 genes) (log2FC > 2; $p < 0.05$).

(E–G) Gene ontology enrichment analyses of common differentially upregulated genes between the two groups (E) Neural retina function, development and morphogenesis, (F) Retinal ganglion cell development, (G) Amacrine cell development. Some of the main GO groups and subgroups are shown with some indicative genes and the respective log2FC in $MKS^{01/02}$ (orange column) and BBS^{01} RSs (green column) ($p < 0.05$).

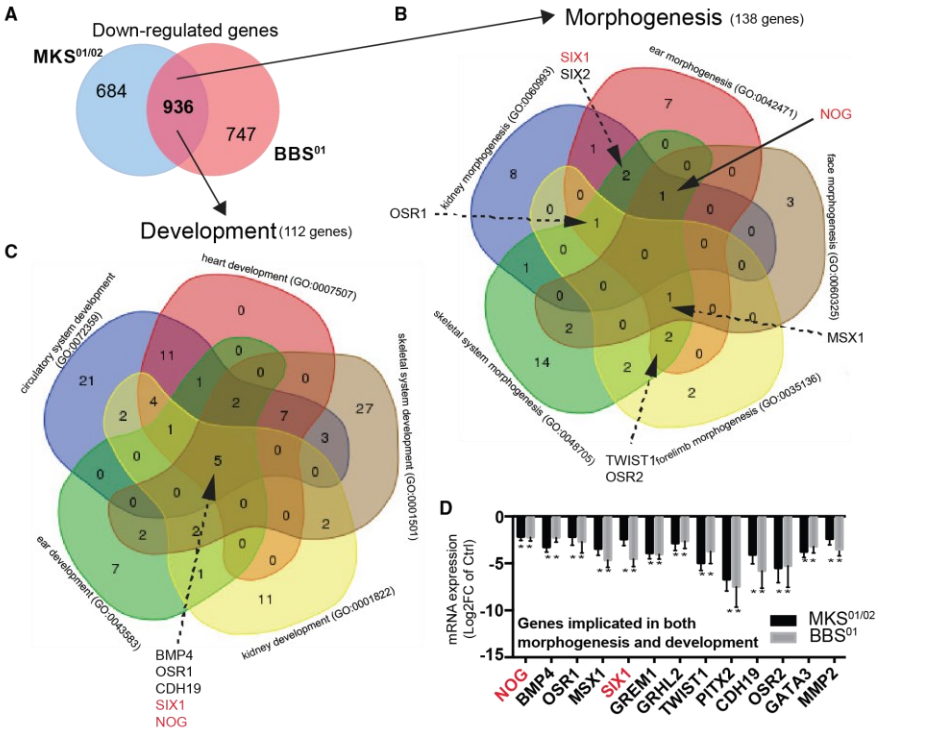


Figure 3. Reduced Expression of Genes Involved in Development and Morphogenesis

(A) Venn diagrams showing the intersection of significant genes differentially downregulated between $MKS^{01/02}$ (blue) and BBS^{01} RSs (red).

(B and C) Venn diagrams showing the intersection of significative downregulated genes belonging to GO groups involved in the morphogenesis (B) or in the development (C) of five organs/systems commonly affected in both MKS and BBS (log2FC > 2; $p < 0.05$).

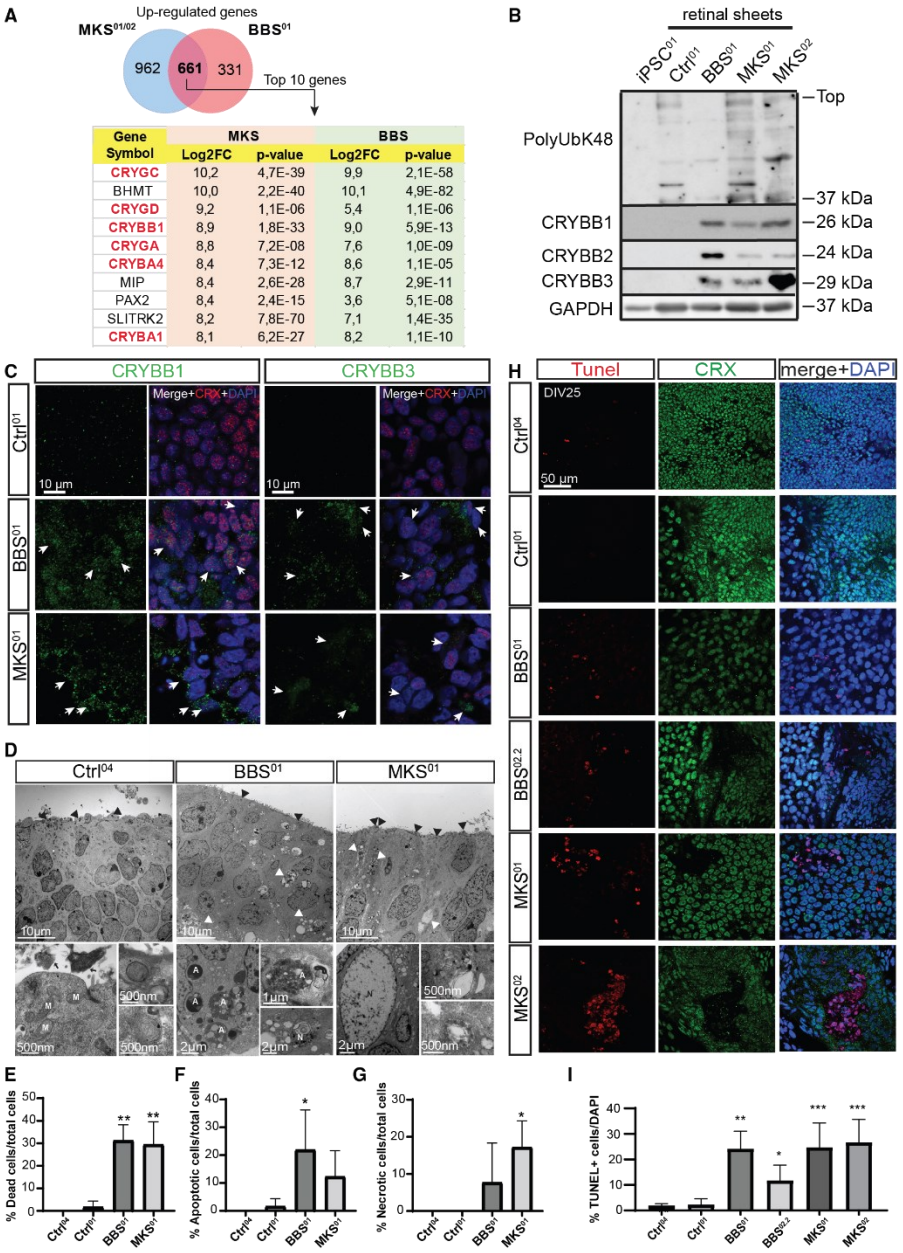
(D) Gene expression levels in BBS^{01} and $MKS^{01/02}$ RSs at DIV60 of genes implicated in both development and morphogenesis. In red, the genes common to all five developmental groups with a role also in morphogenesis.

All values are means ± SEM. *$p < 0.05$, Student's unpaired t test.

S1H). Hence, we could observe by immunoblot the presence of MKS3 in $Ctrl^{01}$, $MKS^{01/02}$, and $BBS^{01/02}$ RSs, and MKS3 levels were even slightly increased in MKS^{01} RSs (Figure 5H). Using antibodies against Centriolin and MKS3, we performed confocal immunofluorescence imaging and 3D reconstruction on dissociated control PRs. We observed three different configurations between Centriolin and MKS3: BBs and daughter centrioles surrounding MKS3 (F′); MKS3 linked to the centriole (F″); and unlinked MKS3 and Centriolin (F‴) (Figure 5F). Notably, while we observed that MKS3 was often in close association with Centriolin in control and BBS^{01} PRs (Pearson's coefficient correlation = 0.29 and 0.23), this association was near absent in MKS^{01} PRs (Pearson's coefficient correlation = 0.06) (Figures 5E–5G). These results suggested defective interaction of MKS3 with the centriole in MKS^{01} PRs, possibly leading to the production of supernumerary centrioles (see model in Figure S6).

MKS^{01} PRs Have Smaller but Supernumerary Cilia and Present Abnormal Accumulations of Ciliary Proteins

Using an antibody against acetylated α-tubulin, we found that supernumerary centrioles in MKS^{01} PRs were also associated with an increasing number of cilia (Figures 6A–6C). This phenotype was not observed in BBS^{01} PRs. We performed quantitative analyses of the images using IMARIS, which revealed that, although more numerous, cilia in MKS^{01} PRs were shorter and thinner than normal (Figures



Gene Symbol	MKS		BBS	
	Log2FC	p-value	Log2FC	p-value
CRYGC	10,2	4,7E-39	9,9	2,1E-58
BHMT	10,0	2,2E-40	10,1	4,9E-82
CRYGD	9,2	1,1E-06	5,4	1,1E-06
CRYBB1	8,9	1,8E-33	9,0	5,9E-13
CRYGA	8,8	7,2E-08	7,6	1,0E-09
CRYBA4	8,4	7,3E-12	8,6	1,1E-05
MIP	8,4	2,6E-28	8,7	2,9E-11
PAX2	8,4	2,4E-15	3,6	5,1E-08
SLITRK2	8,2	7,8E-70	7,1	1,4E-35
CRYBA1	8,1	6,2E-27	8,2	1,1E-10

(legend on next page)

6D and 6E). Some rare giant cilia resembling a fusion of different cilia were also observed. The number of observed cilia was normal in undifferentiated iPSCs from MKS^{01} patients, suggesting cell-type specificity for this phenotype (Figures S5F–S5H). RPGR mostly labels the BBs of the CC (Figure 6F) (He et al., 2008). Using confocal microscopy and 3D reconstruction, we observed that control RSs presented a relatively uniform immuno-labeling for RPGR proximal to the PNA-positive OS (Figure 6G). In contrast, MKS^{01} RSs were disorganized, as revealed by the unequal distribution of PNA. Furthermore, the signal for RPGR was mislocalized and tended to accumulate in aggregates. RPGR was also mislocalized in RSs from the BBS^{01} patient (Figure 6G). These results suggested supernumerary cilia, abnormal retinal tissue morphogenesis, and aggregation of ciliary proteins in MKS PRs.

BBS PRs Display Mitotic Spindle Checkpoint Activation and Genomic Instability

Significantly upregulated genes in BBS^{01} RSs included markers of mitotic spindle checkpoint and regulation of microtubule cytoskeleton organization (Figures 7A, 7B, and S7B), suggesting possible genomic instability during mitosis. On the other hand, we observed downregulation of genes involved in extracellular matrix organization and tube development (Figure S7D). DIV15 PR progenitors were analyzed using specific antibodies. We found that the number of cells expressing the DNA damage-response markers, γH2Ax and 53BP1, and the activated mitotic checkpoint protein, pCHK2, were increased in BBS$^{01/02}$ samples, but not in control or MKS$^{01/02}$ samples (Figures 7C–7G and S7E–S7H). This phenotype was, however, generally more severe in the BBS^{01} case than in the BBS^{02} case. We also observed the presence of very large nuclei in BBS$^{01/02}$ PRs that were also γH2Ax and pCHK2-positive, suggesting possible arrest in the G2/M phase of the cell cycle. γH2Ax, a marker of DNA double-strand breaks, was also detected by immunoblot on whole-cell extracts in BBS^{01} RSs at DIV60 of differentiation, but not in control or MKS^{01} RSs (Figure 7H). Consistently, micronuclei, nuclear bridges, mitotic catastrophe, and nuclei having multipolar or monopolar spindles, were more frequent in BBS^{01} PRs progenitors than in the MKS^{01} or control one. Nuclear bridges were detected in all conditions but tended to be more abundant in BBS^{01} PRs. Mitotic catastrophes were also significantly higher in BBS^{01} than in control cells, but they were also present in MKS^{01} cells (Figure 7I). We concluded that, although MKS and BBS PRs shared a broad number of molecular and cellular alterations, they also displayed unique anomalies that may help explain the distinct phenotypes characterizing these disorders.

DISCUSSION

We generated iPSCs from control and ciliopathy cases and differentiated them into polarized 3D-adherent RSs that could recapitulate normal PR development and disease state, respectively. By WGS, we identified mutations in *BBS10* (BBS^{01} and BBS^{02} patients) and *TMEM67* (MKS^{01} case). Using RNA-seq analyses and high-resolution fluorescence microscopy, we were able to identify and study retinal developmental anomalies and PR degeneration processes that characterized MKS and BBS. Importantly, we could also distinguish specific pathological features that were either common or unique to the two syndromes.

Previous reports based on rare genetic cases suggested that BBS and MKS may represent a "unique disease" with

Figure 4. Induction and Accumulation of Crystallins in MKS and BBS RSs

(A) Venn diagrams showing the intersection of significant genes differentially upregulated between MKS$^{01/02}$ (blue) and BBS^{01} RSs (red) (log2FC > 2; $p < 0.05$), and the top 10 upregulated genes common to the two groups with their respective log2FC and p value for each group. Genes that are members of the crystallin family are highlighted in red.

(B) Immunoblot on extracts from undifferentiated iPSCs and RSs.

(C) Immunofluorescence (IF) representative images showing CRYBB1 and CRYBB3 accumulation in MKS$^{01/02}$ and BBS^{01} CRX-positive cells at DIV45. Note the extracellular and intracellular accumulation of crystallin (white arrows). Scale bars, 10 μm.

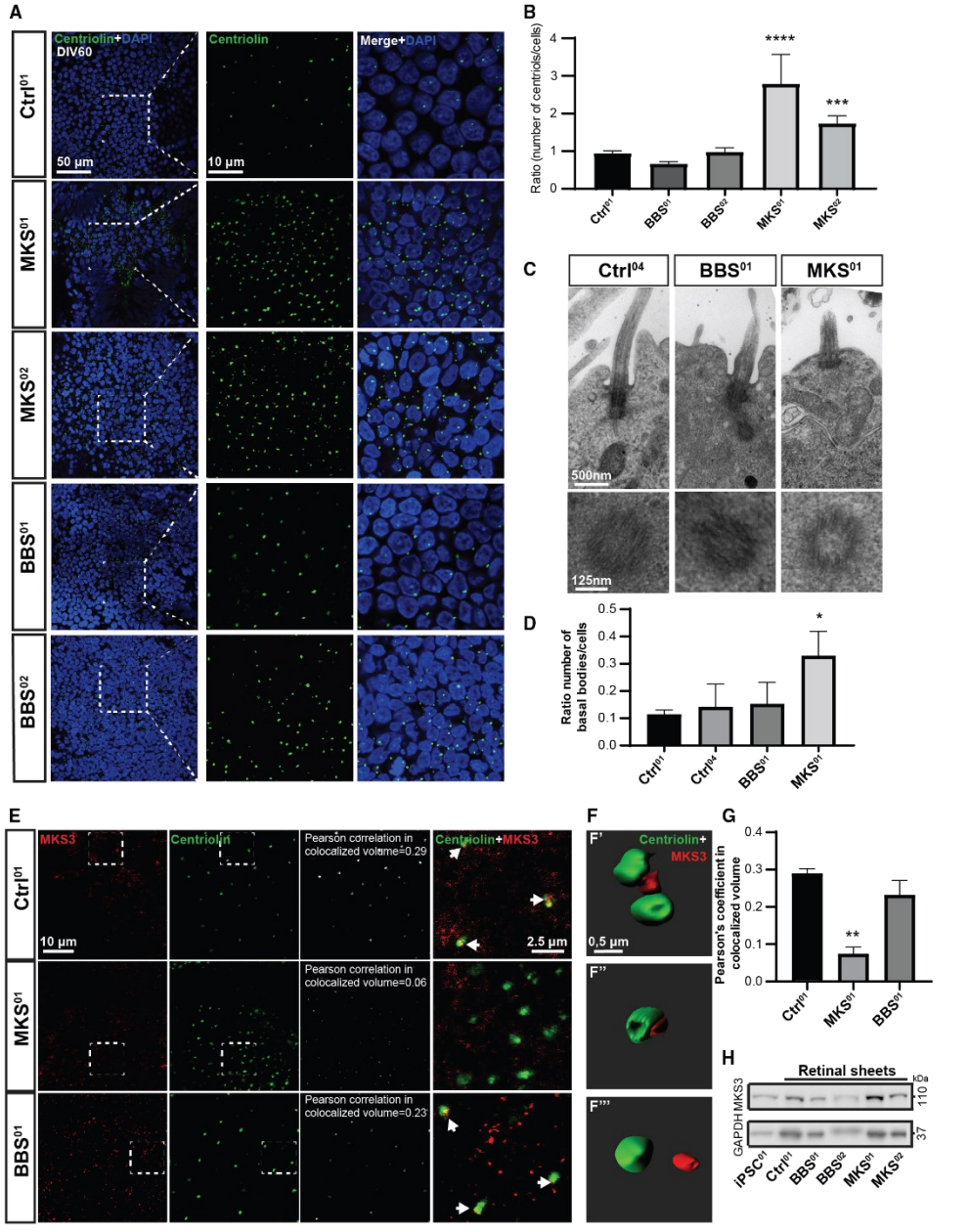
(D) TEM representative images of RSs at DIV60. BBs (black arrowheads) and dead cells (white arrowheads). For each condition at the bottom some high-magnification images. In Ctrl^{04} we can observe IS structure with high density of mitochondria, BBs near the apical surface, and marked gap junction. A beginning of OS with stack-like structures (black arrows) can be observed beyond the apical surface. Vesicles with laminar material are also observed near the surface. Big vesicles similar to lysosomes were particularly abundant in MKS RSs. M, mitochondria; A, apoptotic cells; N, necrotic cells.

(E–G) Quantification of the percentage of total dead cells (E), apoptotic (F), and necrotic (G) cells observed by TEM at DIV60 (n = 3–4 fields for condition, an average of 24 cells for field.

(H) IF representative images in dissociated cells culture at DIV20. Scale bar, 50 μm.

(I) Quantification of the percentage of TUNEL-positive cells over total nuclei at DIV20 (n = 4–5 fields for condition, an average of 85 cells for field).

n = 3 independent biological replicates for all experiments. All values are means ± SEM. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$ by one-way ANOVA test.



(legend on next page)

two different degrees of severity (Brunham, 2009; Leitch et al., 2008). Using our *in vitro* model, we found that 60% of deregulated genes in BBS^{01} PRs and 49% of those in $MKS^{01/02}$ PRs were common between both diseases, although BBS and MKS cases studied here carried mutations in unrelated genes. Brain, eye, and bone malformations as well as polydactyly are common to MKS and BBS. Hence, we found deregulation of hundreds of neurodevelopmental genes, most of them common to both ciliopathies. WNT, NOTCH, and HOX signaling pathways were highly affected, consistent with the essential function of WNT and NOTCH in neural development, and with the role of the *HOX* gene cluster in axial skeleton development and digit formation (Sheth et al., 2012; Tan et al., 2013; Wheway et al., 2013). Unexpectedly, we observed significant downregulation of *NOG*, *BMP*, *SIX1*, *MSX1*, *TWIST*, *OSR1*, and *PITX2*. *NOG* and *BMP* are important for nervous system, muscle, and bone development, while the others are required for the differentiation of intermediate mesoderm-derivative (gonads and kidneys) and limb buds (*OSR1*), in epithelial-mesenchymal transition (*TWIST*), in the development of the neck, ears, and kidneys (*SIX1*), and of oral structures (*MSX1*). *PITX2* is involved in the establishment of the left-right axis and morphogenesis of many organs (Campione et al., 1999; Zacharias et al., 2011). Mutations in *PITX2* are associated with Axenfeld-Rieger syndrome, iridogoniodysgenesis syndrome, and sporadic cases of Peters anomaly, revealing its important role in eye development (Kozlowski and Walter, 2000). Notably, polydactyly, craniofacial, neural, retinal, ocular, skeletal, and kidney anomalies are commonly observed in syndromic ciliopathies (Brunham, 2009; Novarino et al., 2011). Thus, developmental defects observed in peripheral organs and musculoskeletal structures of MKS and BBS patients were revealed in our gene expression profile analyses of RSs.

We found upregulation of several apoptotic genes and genes of the beta- and gamma-crystallin families in both MKS and BBS RSs. Alpha-crystallins are known to operate as molecular chaperones for misfolded proteins and can be found in ubiquitin-associated inclusions in neurodegenerative diseases (Lowe et al., 2001; Thornell and Aquilina, 2015). In contrast, the function of beta- and gamma-crystallins in non-lens tissue is not well understood. Some studies have linked the expression of beta- and gamma-crystallins to retinal diseases. It has also been observed that gamma- and beta-crystallin production increases before PR cell death in animal models of retinitis pigmentosa and light-induced retinal degeneration (Fort et al., 2009; Organisciak et al., 2006; Piri et al., 2007; Sakaguchi et al., 2003). βB2-Crystallin (*CRYBB2*) was shown to be strongly expressed in regenerating ganglion cells where it may promote axonal regrowth (Liedtke et al., 2007). Thus, although the exact role of crystallins in neurodegenerative contexts remains to be clarified, our findings reveal robust activation of this pathway in PRs from ciliopathy patients.

In MKS^{01} PRs, we observed near absent co-localization of MKS3 with Centriolin, which may explain the presence of supernumerary centrioles. Cilia were also shorter and thinner than normal, suggesting cilia degeneration and/or abnormal formation and maintenance. Notably, it was shown that, in mouse renal-tubule epithelial cells, reducing the level of MKS3 using small interfering RNA impaired the number of cilia, whereas its complete loss caused elongated cilia (Cook et al., 2009). Multi-ciliated cells were also observed in MKS patients and in a rat model of MKS3 (Gattone et al., 2004). We also found that MKS^{01} RSs were highly disorganized and that the signal for RPGR was mislocalized and tended to accumulate in aggregates. BBS^{01} cells also showed abnormal RPGR localization. Unlike MKS^{01} PRs, however, RPGR was generally mislocalized at the basement of BBS^{01} PRs, suggesting abnormal transport or docking to the BB. This reveals both similarities and differences between the two syndromes. Moreover, genes implicated in canonical WNT signaling were upregulated, and this was apparently at the expense of the non-canonical one. In polarized cells, the cilium grows

Figure 5. Supernumerary Centrioles in MKS PRs

(A) IF images of PRs at DIV60. Composed image from two focal planes (DAPI and Centriolin). Scale bars, 50 μm (left), 10 μm (right).

(B) Quantification of the number of centrioles (Centriolin) per cells at DIV60 (n = 4 fields for condition, an average of 113 cells for a field).

(C) TEM representative images of cilia and BBs and its section (bottom).

(D) Quantification of the ratio of BBs/cell per field by TEM. N = 3–4 images for condition an average of 24 cells for a field.

(E) IF images of dissociated PRs at DIV60. The white channel represents colocalized volume between Centriolin and MKS3 with the respective Pearson correlation value quantified in (G). On the right, high-magnification images of the respective dashed square in each condition.

(F) Representative surface rendering from 3D z stack reconstruction of the three conformations observed between Centriolin and MKS3: BBs and daughter centrioles surrounding MKS3 (F′), MKS3 linked to Centriolin (F″), and MKS3 and Centriolin unbound (F‴). Scale bar, 0.5 μm.

(G) Pearson correlation of colocalized volume between Centriolin and MKS3 related to (E).

(H) Immunoblot on extracts from iPSCs and RSs.

n = 3 independent biological replicates for all experiments. All values are means ± SEM. *p < 0.05, **p < 0.01, ***p < 0.001, ****p < 0.0001 by one-way ANOVA test.

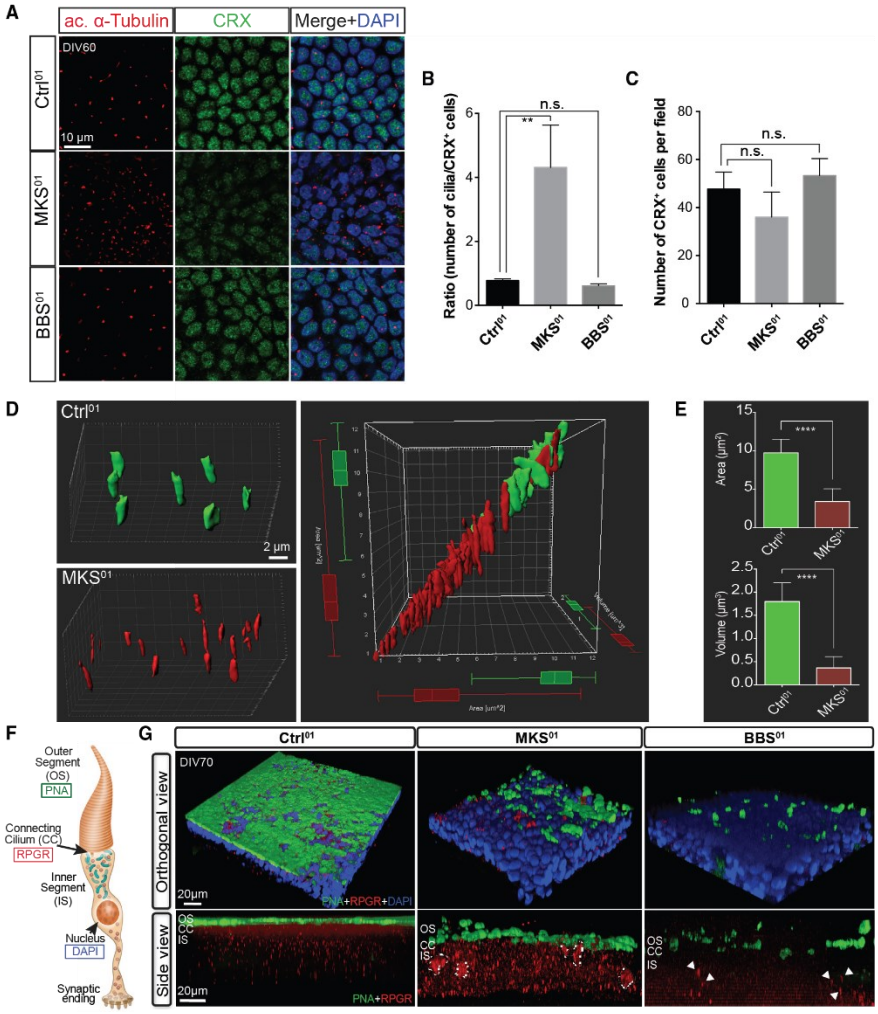


Figure 6. Supernumerary Cilia and Aggregation of Ciliary Proteins in MKS^{01} PRs

(A) IF of PRs at DIV60. Composed image from two focal planes (DAPI and acetyl-α-tubulin). Scale bar, 10 μm.

(B and C) Quantification of the number of cilia (acetyl-α-tubulin) per cells (B) and the number of CRX-positive cells per field (C) in cells at DIV60 (n = 4 fields for condition, an average of 53 cells for a field).

(D) Surface rendering 3D reconstruction from confocal z stack images of Ctrl^{01} (green) and MKS^{01} (red) based on acetyl-α-tubulin staining (left). Scale bar, 2 μm. A 3D plot of cilia (cilia from three images put together for each condition) based on area and volume (right). Notice how control cilia clustered together while MKS^{01} cilia are more heterogeneous, smaller, and more abundant (n = 2 independent experiments).

(E) Quantification of area and volume of Ctrl^{01} and MKS^{01} cilia (ctrl n = 20; MKS n = 47).

(F) Representation of cone PRs with reference to OS, CC, IS, and nucleus, and the localization of the markers used in (G).

(legend continued on next page)

at the apical side. During ciliogenesis, planar cell polarity and apical-basal polarity are necessary for docking of the centrioles at the plasma membrane, and for correct formation and maintenance of PR BBs (Kim et al., 2010; Oishi et al., 2006). Taken as a whole, these observations suggest that hallmarks of the MKS phenotype previously reported in MKS patients and animal models were recapitulated at the cellular and molecular levels in MKS$^{01/02}$ RSs. Furthermore, we described for the first time the presence of supernumerary cilia and centrioles in MKS$^{01/02}$ PRs, which was associated with reduced co-localization of MKS3 with Centriolin in the MKS^{01} case, suggesting deficient interaction of the mutant MKS3 protein with centrioles. In the MKS^{02} case, which does not carry mutations in *TMEM67*, the number of centrioles/cell was about 2-fold higher than in controls and BBS$^{01/02}$ cases but was less severe than in MKS^{01}. It is not surprising thus that despite carrying mutations in different genes, MKS cases show highly related phenotypes at the cellular and molecular levels.

Previous reports have described cilia proteins with cilia-independent functions, including mitotic spindle generation and mitotic process regulation (Hua and Ferland, 2018; Vertii et al., 2015; Yuan and Sun, 2013). Unexpectedly, we found genomic instability in BBS$^{01/02}$ retinal progenitors and PRs, with the most prominent feature being the activation of the mitotic spindle checkpoint. This suggests a new and essential role of BBS10 in microtubule cytoskeleton organization during mitosis.

In summary, we have produced iPSCs from control, and MKS and BBS cases, and differentiated them into RSs, allowing recapitulation of normal and pathological human retinal and PRs development *in vitro*. Molecular and cell biological analyses further revealed known and novel disease mechanisms associated with retinal ciliopathies, opening potentially new avenues for disease treatment.

EXPERIMENTAL PROCEDURES

iPSC Generation and Characterization

Human PSCs were used in accordance with the Canadian Institute Health Research (CIHR) guidelines and approved by the "Comité de Surveillance de la Recherche sur les Cellules Souches" of the CIHR and the Maisonneuve-Rosemont Hospital Ethic Committee. MKS and BBS fibroblasts were obtained from clinically diagnosed individuals (Coriell Biorepository). These were reprogrammed with the Yamanaka factors, *OCT4*, *SOX2*, *KLF4*, and *c-MYC*, using the pMIG vector set. Stem cells were grown on MEF feeder layers (Global Stem Cell GSC-6001G) in iPSC medium. To generate teratomas, ~3 × 10^6 undifferentiated iPSCs were implanted beneath the neck scruff of non-obese diabetic-severe combined immunodeficiency immunodeficient mice. Tumors were harvested 12 weeks later and processed for histology analysis. See Supplemental Information for details.

Differentiation of Human ESCs and iPSCs into RSs

iPSCs and ESCs were dissociated using ReLeSR (STRMCELL Technologies, cat. no. 05,872) and platted on growth factor reduced Matrigel (Corning, no. 356231) in StemFlex cell medium (Gibco, no. A3349401) supplemented with ROCK-inh (Y-27632; 10 μM, Cayman Chemical, no. 10005583). When PSCs reach full confluence, the medium was switched to differentiation medium supplemented with 60 ng/mL of recombinant COCO as described in Zhou et al. (2015). See Supplemental Information for details.

RNA-Seq and WGS Analyses

RNA-seq libraries were prepared using Ion Total RNA-Seq Kit v.2. Libraries were sequenced onto P1 chips from Ion Torrent as unpaired to reach 40 million reads for each sample. Raw sequencing files (FASTQ) were validated using FASTQC v.0.11.7.

Base calling for WGS was performed using Illumina HiSeq Analysis Software (v.2—2.5.55.1311). Reads were mapped to the b37 reference sequence. See Supplemental Information for details.

Statistical Analysis

Statistical analysis was performed using GraphPad software (Prism 6). Statistical differences were analyzed using the Student's t test for unpaired samples. Values are representative of at least three experiments. When comparisons were made using independent samples of equal size and variance following a normal distribution, significance was assessed using an unpaired two-sided Student's t test. Where several groups were compared, significance was assessed by ANOVA and adjusted for multiple comparisons using the Bonferroni correction. Differential expression was assayed using a log2 fold change statistical algorithm or one-way ANOVA with a p value cutoff at 0.05. For gene ontology, a false discovery rate cutoff of 0.01 was applied.

ACCESSION NUMBERS

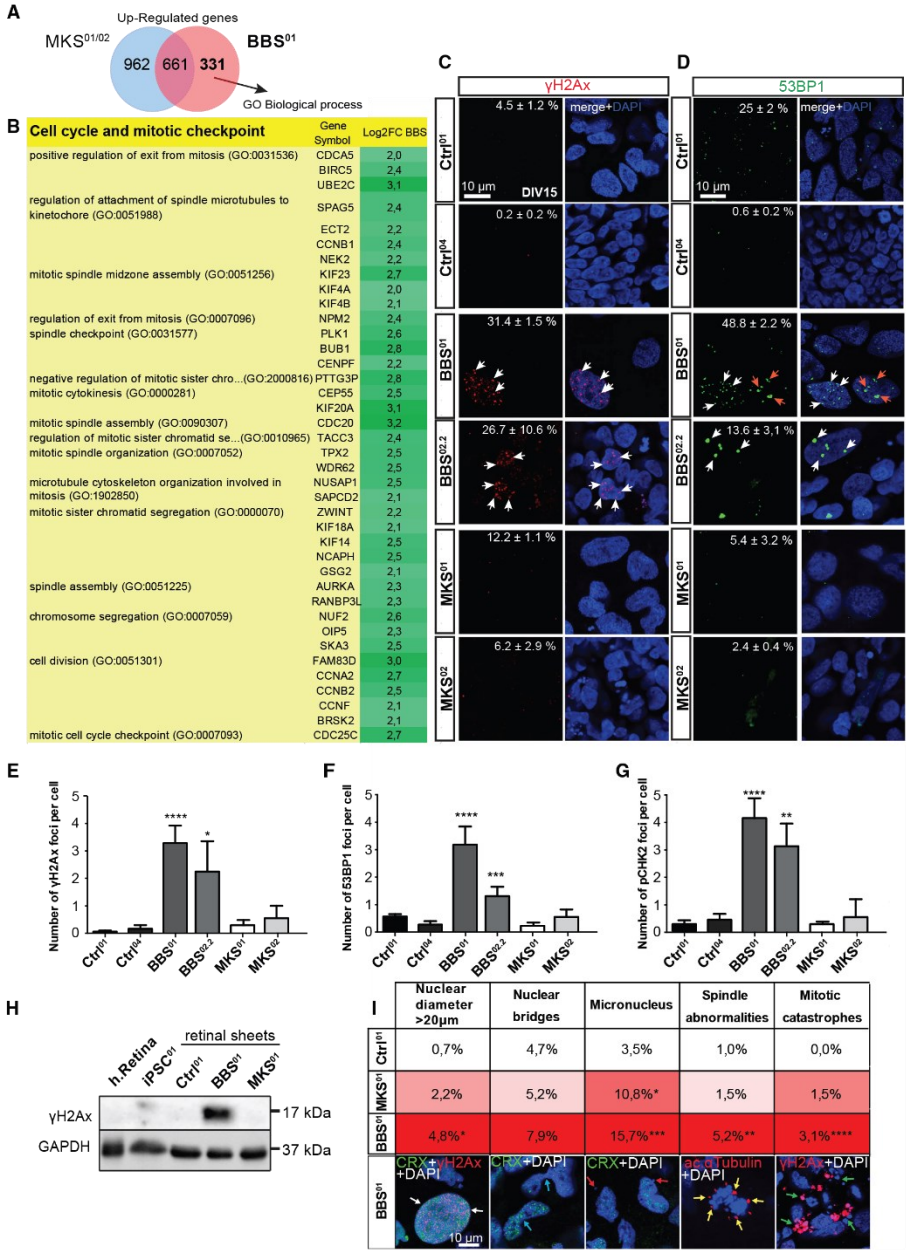
RNAseq Data was deposited in the GEO: GSE133247. WGS data were deposited in BioSample database: SAMN13656497; SAMN13656498

SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at https://doi.org/10.1016/j.stemcr.2020.02.005.

(G) 3D z stack reconstruction of RSs from IF images at DIV70. Dotted lines, intracellular aggregates of RPGR. White arrows, RPGR not polarized to the apical zone. Scale bars, 20 μm.

n = 3 independent biological replicates for all experiments. All values are means ± SEM. **p < 0.01, ****p < 0.0001, Student's unpaired t test.



(legend on next page)

AUTHOR CONTRIBUTIONS

A.B. and G.B. wrote the article. A.B. did the cell culture and most of the experiments and analyses. B.S.F. generated and characterized MKS and BBS iPSc lines and helped write the manuscript. E.H. helped identify the mutations through WGS and helped write the manuscript. A.F. contributed to the bioinformatic analysis of the RNA-seq and helped write the manuscript. R.H. contributed to the bioinformatic analysis of the RNA-seq and to the quantifications of DNA damage and helped write the manuscript.

ACKNOWLEDGMENTS

This work was supported by grants from the Foundation Fighting Blindness Canada, Natural Sciences and Engineering Research Council of Canada (2017-05504), and Maisonneuve-Rosemont Hospital Foundation in Montreal, Canada. A.F., R.H., and A.B. were supported by fellowships from the Molecular Biology Program of Université de Montréal. Work in the Freedman laboratory was supported by NIH Award K01DK102826, United States. We thank Steve Breault, Anna Baccei, and Joseph Bonventre for assistance with scanning electron microscopy and stem cell culture, Eric Pierce, Marc Consugar, and Erika Tavares for the mutational analysis and WGS, and Fabien Martinel for the photoreceptor graphics. G.B. and A.F. are co-founders of StemAxon and members of its scientific advisory board. The corporation was not involved in this study.

Received: July 8, 2019
Revised: February 10, 2020
Accepted: February 11, 2020
Published: March 10, 2020

REFERENCES

Aboshiha, J., Dubis, A.M., Carroll, J., Hardcastle, A.J., and Michaelides, M. (2016). The cone dysfunction syndromes. Br. J. Ophthalmol. *100*, 115–121.

Adams, M., Simms, R.J., Abdelhamed, Z., Dawe, H.R., Szymanska, K., Logan, C.V., Wheway, G., Pitt, E., Gull, K., Knowles, M.A., et al. (2012). A meckelin-filamin A interaction mediates ciliogenesis. Hum. Mol. Genet. *21*, 1272–1286.

Adams, N.A., Awadein, A., and Toma, H.S. (2007). The retinal ciliopathies. Ophthal. Genet. *28*, 113–125.

Ahmed, M., Xu, J., and Xu, P.-X. (2012). EYA1 and SIX1 drive the neuronal developmental program in cooperation with the SWI/SNF chromatin-remodeling complex and SOX2 in the mammalian inner ear. Development *139*, 1965–1977.

Álvarez-Satta, M., Castro-Sánchez, S., and Valverde, D. (2017). Bardet-Biedl syndrome as a chaperonopathy: dissecting the major role of chaperonin-like BBS proteins (BBS6-BBS10-BBS12). Front. Mol. Biosci. *4*, 1–7.

Breuer, D.K., Yashar, B.M., Filippova, E., Hiriyanna, S., Lyons, R.H., Mears, A.J., Asaye, B., Acar, C., Vervoort, R., Wright, A.F., et al. (2002). A comprehensive mutation analysis of RP2 and RPGR in a North American cohort of families with X-linked retinitis pigmentosa. Am. J. Hum. Genet. *70*, 1545–1554.

Brunham, L. (2009). Cilia get serious: Meckel-Gruber and Bardet-Biedl syndromes represent a spectrum of allelic disorders. Clin. Genet. *75*, 40–41.

Burnicka-Turek, O., Steimle, J.D., Huang, W., Felker, L., Kamp, A., Kweon, J., Peterson, M., Reeves, R.H., Maslen, C.L., Gruber, P.J., et al. (2016). Cilia gene mutations cause atrioventricular septal defects by multiple mechanisms. Hum. Mol. Genet. *25*, 3011–3028.

Campione, M., Steinbeisser, H., Schweickert, A., Deissler, K., van Bebber, F., Lowe, L.A., Nowotschin, S., Viebahn, C., Haffter, P., Kuehn, M.R., et al. (1999). The homeobox gene Pitx2: mediator of asymmetric left-right signaling in vertebrate heart and gut looping. Development *126*, 1225–1234.

Chang, P., Giddings, T.H., Winey, M., and Stearns, T. (2003). ϵ-Tubulin is required for centriole duplication and microtubule organization. Nat. Cell Biol. *5*, 71–76.

Cook, S.A., Collin, G.B., Bronson, R.T., Naggert, J.K., Liu, D.P., Akeson, E.C., and Davisson, M.T. (2009). A mouse model for Meckel syndrome type 3. J. Am. Soc. Nephrol. *20*, 753–764.

Delous, M., Baala, L., Salomon, R., Laclef, C., Vierkotten, J., Tory, K., Golzio, C., Lacoste, T., Besse, L., Ozilou, C., et al. (2007). The ciliary gene RPGRIP1L is mutated in cerebello-oculo-renal syndrome (Joubert syndrome type B) and Meckel syndrome. Nat. Genet. *39*, 875–881.

Figure 7. Activation of the Mitotic Spindle Checkpoint and Accumulation of DNA Damage in BBS PRs

(A) Venn diagrams showing the intersection of significant genes differentially upregulated in $MKS^{01/02}$ (blue) and BBS^{01} RSs (red) (log2FC > 2; p < 0.05).

(B) Gene ontology enrichment analyses of BBS^{01}-specific differentially upregulated genes. The main GO group and subgroups are shown. Some representative genes and the respective log2FC (p < 0.05) are indicated for each category.

(C and D) IF images of PR progenitors at DIV15 for γH2Ax (C), and 53BP1 (D). The percentage of positive cells (more than three foci per cell) are indicated on the respective images (average n = 253 cells for each condition). Scale bars, 10 μm.

(E–G) Quantification of the average number of foci per cell for γH2Ax (E), 53BP1 (F), and pCHK2 (G) in PR progenitors at DIV15 (n > 3 fields for condition, an average of 84 cells for a field).

(H) Immunoblot on extracts from undifferentiated iPSCs and RSs at DIV60.

(I) Representative images for common mitotic abnormalities and the respective percentages of cells for each condition in PR progenitors at DIV15. White arrows, giant nucleus; blue arrows, nuclear bridge; red arrows, micronuclei; yellow arrows, multifocal spindle; green arrows, fragmented nuclei strongly marked by the anti-γH2Ax antibody. Scale bars, 10 μm.

n = 3 independent biological replicates for all experiments. All values are means ± SEM. *p < 0.05, **p < 0.01, ***p < 0.001, ****p < 0.0001, Student's unpaired t test.

Fort, P.E., Freeman, W.M., Losiewicz, M.K., Singh, R.S.J., and Gardner, T.W. (2009). The retinal proteome in experimental diabetic retinopathy: up-regulation of crystallins and reversal by systemic and periocular insulin. Mol. Cell. Proteomics *8*, 767–779.

Gattone, V.H., Tourkow, B.A., Trambaugh, C.M., Yu, A.C., Whelan, S., Phillips, C.L., Harris, P.C., and Peterson, R.G. (2004). Development of multiorgan pathology in the wpk rat model of polycystic kidney disease. Anat. Rec. *277A*, 384–395.

Goodnough, L.H., Dinuoscio, G.J., and Atit, R.P. (2016). Twist1 contributes to cranial bone initiation and dermal condensation by maintaining Wnt signaling responsiveness. Dev. Dyn. *245*, 144–156.

Grandy, R., Tomaz, R.A., and Vallier, L. (2019). Modeling disease with human inducible pluripotent stem cells. Annu. Rev. Pathol. Mech. Dis. *14*, 449–468.

Gupta, G.D., Coyaud, É., Gonçalves, J., Mojarad, B.A., Liu, Y., Wu, Q., Gheiratmand, L., Comartin, D., Tkach, J.M., Cheung, S.W.T., et al. (2015). A dynamic protein interaction landscape of the human centrosome-cilium interface. Cell *163*, 1484–1499.

He, S., Parapuram, S.K., Hurd, T.W., Behnam, B., Margolis, B., Swaroop, A., and Khanna, H. (2008). Retinitis pigmentosa GTPase regulator (RPGR) protein isoforms in mammalian retina: insights into X-linked Retinitis Pigmentosa and associated ciliopathies. Vis. Res. *48*, 366–376.

Hoshino, A., Ratnapriya, R., Brooks, M.J., Chaitankar, V., Wilken, M.S., Zhang, C., Starostik, M.R., Gieser, L., La Torre, A., Nishio, M., et al. (2017). Molecular anatomy of the developing human retina. Dev. Cell *43*, 763–779.e4.

Hua, K., and Ferland, R.J. (2018). Primary cilia proteins: ciliary and extraciliary sites and functions. Cell. Mol. Life Sci. *75*, 1521–1540.

HUGO (2019). Cilia and flagella associated (CFAP) gene family. HUGO Gene Nomenclature Committee. URL: https://www.genenames.org/data/genegroup/#!/group/1491

Kamachi, Y., Uchikawa, M., Tanouchi, A., Sekido, R., and Kondoh, H. (2001). Pax6 and SOX2 form a co-DNA-binding partner complex that regulates initiation of lens development. Genes Dev. *15*, 1272–1286.

Khanna, H. (2015). Photoreceptor sensory cilium: traversing the ciliary gate. Cells *4*, 674–686.

Kim, S.K., Shindo, A., Park, T.J., Oh, E.C., Ghosh, S., Gray, R.S., Lewis, R.A., Johnson, C.A., Attie-Bittach, T., Katsanis, N., et al. (2010). Planar cell polarity acts through septins to control collective cell movement and ciliogenesis. Science *329*, 1337–1340.

Kozlowski, K., and Walter, M.A. (2000). Variation in residual PITX2 activity underlies the phenotypic spectrum of anterior segment developmental disorders. Hum. Mol. Genet. *9*, 2131–2139.

Kyttälä, M., Tallila, J., Salonen, R., Kopra, O., Kohlschmidt, N., Paavola-Sakki, P., Peltonen, L., and Kestilä, M. (2006). MKS1, encoding a component of the flagellar apparatus basal body proteome, is mutated in Meckel syndrome. Nat. Genet. *38*, 155–157.

Leitch, C.C., Zaghloul, N.A., Davis, E.E., Stoetzel, C., Diaz-Font, A., Rix, S., Alfadhel, M., Lewis, R.A., Eyaid, W., Banin, E., et al. (2008). Hypomorphic mutations in syndromic encephalocele genes are associated with Bardet-Biedl syndrome. Nat. Genet. *40*, 443–448.

Li, M., Jia, C., Kazmierkiewicz, K.L., Bowman, A.S., Tian, L., Liu, Y., Gupta, N.A., Gudiseva, H.V., Yee, S.S., Kim, M., et al. (2014). Comprehensive analysis of gene expression in human retina and supporting tissues. Hum. Mol. Genet. *23*, 4001–4014.

Li, X., Ohgi, K.A., Zhang, J., Krones, A., Bush, K.T., Glass, C.K., Nigam, S.K., Aggarwal, A.K., Maas, R., Rose, D.W., et al. (2003). Eya protein phosphatase activity regulates Six1-Dach-Eya transcriptional effects in mammalian organogenesis. Nature *426*, 247–254.

Liedtke, T., Schwamborn, J.C., Schröer, U., and Thanos, S. (2007). Elongation of axons during regeneration involves retinal crystallin beta b2 (crybb2). Mol. Cell. Proteomics *6*, 895–907.

Loncarek, J., and Bettencourt-Dias, M. (2018). Building the right centriole for each cell type. J. Cell Biol. *217*, 823–835.

Lowe, J., Mayer, J., Landon, M., and Layfield, R. (2001). Ubiquitin and the molecular pathology of neurodegenerative diseases. Adv. Exp. Med. Biol. *487*, 169–186.

Ma, L., Wang, Y., Hui, Y., Du, Y., Chen, Z., Feng, H., Zhang, S., Li, N., Song, J., Fang, Y., et al. (2019). WNT/NOTCH pathway is essential for the maintenance and expansion of human MGE progenitors. Stem Cell Reports *12*, 934–949.

Marszalek, J.R., Liu, X., Roberts, E.A., Chui, D., Marth, J.D., Williams, D.S., and Goldstein, L.S.B. (2000). Genetic evidence for selective transport of opsin and arrestin by kinesin-II in mammalian photoreceptors. Cell *102*, 175–187.

May-Simera, H.L., and Kelley, M.W. (2012). Cilia, Wnt signaling, and the cytoskeleton. Cilia *1*, 7.

Navarro Quiroz, E., Navarro Quiroz, R., Ahmad, M., Gomez Escorcia, L., Villarreal, J.L., Fernandez Ponce, C., and Aroca Martinez, G. (2018). Cell signaling in neuronal stem cells. Cells *7*. https://doi.org/10.3390/cells7070075.

Nishimura, D.Y., Fath, M., Mullins, R.F., Searby, C., Andrews, M., Davis, R., Andorf, J.L., Mykytyn, K., Swiderski, R.E., Yang, B., et al. (2004). Bbs2-null mice have neurosensory deficits, a defect in social dominance, and retinopathy associated with mislocalization of rhodopsin. Proc. Natl. Acad. Sci. U S A *101*, 16588–16593.

Nonaka, S., Tanaka, Y., Okada, Y., Takeda, S., Harada, A., Kanai, Y., Kido, M., and Hirokawa, N. (1998). Randomization of left-right asymmetry due to loss of nodal cilia generating leftward flow of extraembryonic fluid in mice lacking KIF3B motor protein. Cell *95*, 829–837.

Novarino, G., Akizu, N., and Gleeson, J.G. (2011). Modeling human disease in humans: the ciliopathies. Cell *147*, 70–79.

Oishi, I., Kawakami, Y., Raya, Á., Callol-Massot, C., and Belmonte, J.C.I. (2006). Regulation of primary cilia formation and left-right patterning in zebrafish by a noncanonical Wnt signaling mediator, duboraya. Nat. Genet. *38*, 1316–1322.

OMIM (2019a). OMIM Entry Search—MECKEL-GRUBER SYNDROME 1 # 249000. URL: https://omim.org/entry/249000

OMIM (2019b). OMIM Entry Search—BARDET-BIEDL SYNDROME 1 # 209900. URL: https://omim.org/entry/209900

Organisciak, D., Darrow, R., Gu, X., Barsalou, L., and Crabb, J.W. (2006). Genetic, age and light mediated effects on crystallin protein expression in the retina. Photochem. Photobiol. *82*, 1088.

Ou, Y.Y., Mack, G.J., Zhang, M., and Rattner, J.B. (2002). CEP110 and ninein are located in a specific domain of the centrosome associated with centrosome maturation. J. Cell Sci. *115*, 1825–1835.

Paradowska-Stolarz, A. (2015). MSX1 gene in the etiology orofacial deformities. Postepy Hig. Med. Dosw. (Online) *69*, 1499–1504.

Pazour, G.J., Baker, S.A., Deane, J.A., Cole, D.G., Dickert, B.L., Rosenbaum, J.L., Witman, G.B., and Besharse, J.C. (2002). The intraflagellar transport protein, IFT88, is essential for vertebrate photoreceptor assembly and maintenance. J. Cell Biol. *157*, 103–114.

Piri, N., Song, M., Kwong, J.M.K., and Caprioli, J. (2007). Modulation of alpha and beta crystallin expression in rat retinas with ocular hypertension-induced ganglion cell degeneration. Brain Res. *1141*, 1–9.

Reiter, J.F., Blacque, O.E., and Leroux, M.R. (2012). The base of the cilium: roles for transition fibres and the transition zone in ciliary formation, maintenance and compartmentalization. EMBO Rep. *13*, 608–618.

Sahly, I., Dufour, E., Schietroma, C., Michel, V., Bahloul, A., Perfettini, I., Pepermans, E., Estivalet, A., Carette, D., Aghaie, A., et al. (2012). Localization of Usher 1 proteins to the photoreceptor calyceal processes, which are absent from mice. J. Cell Biol. *199*, 381–399.

Sakaguchi, H., Miyagi, M., Darrow, R.M., Crabb, J.S., Hollyfield, J.G., Organisciak, D.T., and Crabb, J.W. (2003). Intense light exposure changes the crystallin content in retina. Exp. Eye Res. *76*, 131–133.

Satir, P., Guerra, C., and Bell, A.J. (2007). Evolution and persistence of the cilium. Cell Motil. Cytoskeleton *64*, 906–913.

Sheth, R., Marcon, L., Bastida, M.F., Junco, M., Quintana, L., Dahn, R., Kmita, M., Sharpe, J., and Ros, M.A. (2012). Hox genes regulate digit patterning by controlling the wavelength of a Turing-type mechanism. Science *338*, 1476–1480.

Smith, H.L., Li, W., and Cheetham, M.E. (2015). Molecular chaperones and neuronal proteostasis. Semin. Cell Dev. Biol. *40*, 142–152.

Smith, U.M., Consugar, M., Tee, L.J., McKee, B.M., Maina, E.N., Whelan, S., Morgan, N.V., Goranson, E., Gissen, P., Lilliquist, S., et al. (2006). The transmembrane protein meckelin (MKS3) is mutated in Meckel-Gruber syndrome and the wpk rat. Nat. Genet. *38*, 191–196.

Stoetzel, C., Laurier, V., Davis, E.E., Muller, J., Rix, S., Badano, J.L., Leitch, C.C., Salem, N., Chouery, E., Corbani, S., et al. (2006). BBS10 encodes a vertebrate-specific chaperonin-like protein and is a major BBS locus. Nat. Genet. *38*, 521–524.

Tan, F.E., Vladar, E.K., Ma, L., Fuentealba, L.C., Hoh, R., Espinoza, F.H., Axelrod, J.D., Alvarez-Buylla, A., Stearns, T., Kintner, C., et al. (2013). Myb promotes centriole amplification and later steps of the multiciliogenesis program. Development *140*, 4277–4286.

Thornell, E., and Aquilina, A. (2015). Regulation of αA- and αB-crystallins via phosphorylation in cellular homeostasis. Cell. Mol. Life Sci. *72*, 4127–4137.

Tzekov, R., Stein, L., and Kaushal, S. (2011). Protein misfolding and retinal degeneration. Cold Spring Harb. Perspect. Biol. *3*, a007492.

Vertii, A., Bright, A., Delaval, B., Hehnly, H., and Doxsey, S. (2015). New frontiers: discovering cilia-independent functions of cilia proteins. EMBO Rep. *16*, 1275–1287.

Wheway, G., Abdelhamed, Z., Natarajan, S., Toomes, C., Inglehearn, C., and Johnson, C.A. (2013). Aberrant Wnt signalling and cellular over-proliferation in a novel mouse model of Meckel-Gruber syndrome. Dev. Biol. *377*, 55–66.

Wheway, G., Nazlamova, L., and Hancock, J.T. (2018). Signaling through the primary cilium. Front. Cell Dev. Biol. *6*, 1–13.

Wu, W., Huang, R., Wu, Q., Li, P., Chen, J., Li, B., and Liu, H. (2014). The role of Six1 in the genesis of muscle cell and skeletal muscle development. Int. J. Biol. Sci. *10*, 983–989.

Young, R.W. (1968). Passage of newly formed protein through the connecting cilium of retina rods in the frog. J. Ultrastruct. Res. *23*, 462–473.

Yuan, S., and Sun, Z. (2013). Expanding horizons: ciliary proteins reach beyond cilia. Annu. Rev. Genet. *47*, 353–376.

Zacharias, A.L., Lewandoski, M., Rudnicki, M.A., and Gage, P.J. (2011). Pitx2 is an upstream activator of extraocular myogenesis and survival. Dev. Biol. *349*, 395–405.

Zhou, S., Flamier, A., Abdouh, M., Tétreault, N., Barabino, A., Wadhwa, S., and Bernier, G. (2015). Differentiation of human embryonic stem cells into cone photoreceptors through simultaneous inhibition of BMP, TGFβ and Wnt signaling. Development *142*, 3294–3306.

Stem Cell Reports, Volume *14*

Supplemental Information

Deregulation of Neuro-Developmental Genes and Primary Cilium Cytoskeleton Anomalies in iPSC Retinal Sheets from Human Syndromic Ciliopathies

Andrea Barabino, Anthony Flamier, Roy Hanna, Elise Héon, Benjamin S. Freedman, and Gilbert Bernier

Supplemental Information

Deregulation of neuro-developmental genes and primary cilium cytoskeleton anomalies in iPSC retinal sheets from human syndromic ciliopathies

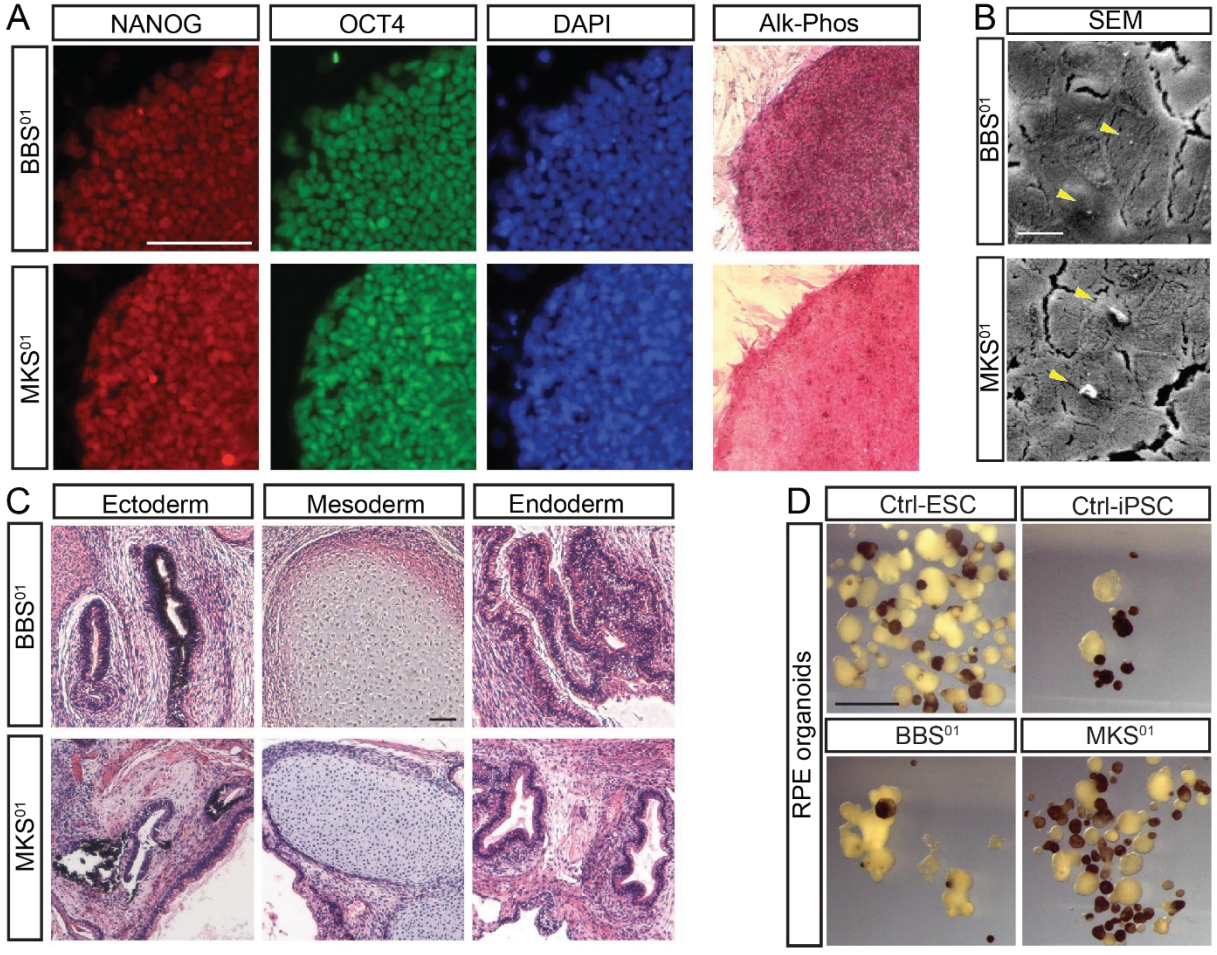
Andrea Barabino[1], Anthony Flamier[2], Roy Hanna[1], Elise Héon[3], Benjamin Freedman[4*], and Gilbert Bernier[1, 5, 6 **]

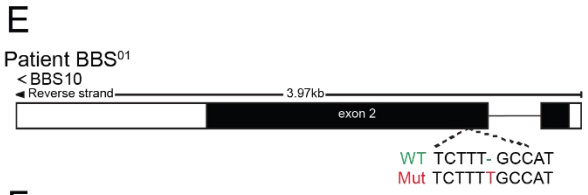
Content:

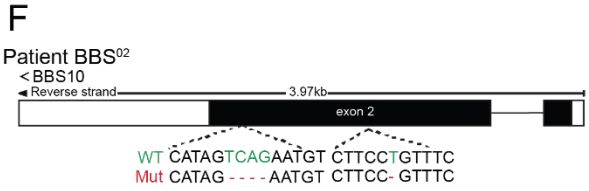
7 Supplementary Figures and legend

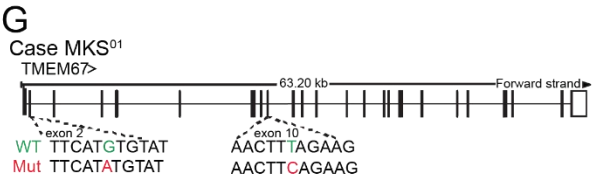
Supplementary Material and Method and tables

Word characters: 18 333 (whole document)









H

Patient identifier	Coriell ID	Coriell Family	Gene	NCBI Reference Sequence:	cChange	pChange
BBS01	GM05948	891	BBS10	NM_024685.4	c.271dup	p.Cys91Leufs*5
					c.271dup	p.Cys91Leufs*5
BBS02	GM05950	892	BBS10	NM_024685.4	c.909_912del	p.Ser303Argfs*3
					c.687del	p.Val230Phefs*7
MKS01	GM07214	-	TMEM67	NM_153704.5	c.233G>A	p.Cys78Tyr
					c.1046T>C	p.Leu349Ser

S. Figure 1

Barabino at al. 2020

A

Developmental process (GO:0032502)	Gene Symbol	Log2FC MKS	Log2FC BBS
Homeobox genes	DLX1	6,7	4,6
	VSX1	5,5	4,6
	SIX6	5,3	5,8
	DLX6	5,0	2,2
	DLX2	4,9	3,4
	LHX4	4,8	4,7
	BARHL2	4,6	5,9
	VSX2	4,2	5,2
	DBX1	3,9	4,0
	NKX6-1	3,5	3,9
	CUX2	3,2	4,2
	LHX9	3,2	2,8
	DRGX	3,1	3,6
	LHX3	2,1	2,9
HOX Family	HOXB9	7,5	3,7
	HOXB8	7,4	4,3
	PHOX2A	2,1	3,9

B

Wnt signaling pathway (GO:0016055)	Gene Symbol	Log2FC MKS	Log2FC BBS
positive regulation of canonical Wnt signaling pathway(GO:0090263)	SFRP2	3,4	4,4
	WNT7A	3,2	2,8
	SFRP2	3,4	4,4
canonical Wnt signaling pathway(GO:0060070)	FZD5	4,3	3,6
	FGF8	4,9	3,4
	NDP	5,3	3,2
Wnt signaling pathway, planar cell polarity pathway (GO:0060071)	CTHRC1	-2,8	-3,1
	WNT11	-2,1	-1,3
	FZD6	-1,5	-1,7
	FZD1	-0,9	-0,7
	FZD7	-1,1	-0,9

C

Notch signaling pathway (GO:0007219)	Gene Symbol	Log2FC MKS	Log2FC BBS
	NEUROD4	5,4	4,9
	HES5	4,6	5,6
	DLL4	4,6	5,8
	ONECUT1	4,1	5,0
	MFNG	3,8	5,0
	DLL3	3,4	4,3
	ASCL1	3,3	3,9
	TBX2	2,9	4,8
	GATA2	2,6	3,7
	HES7	2,6	2,9
	DTX4	2,1	2,5
	CNTN1	-2,6	-2,4
	KRT19	-3,5	-2,3
	GSX2	-3,9	-2,4
	KCNA5	-3,1	-2,4
	SNAI2	-3,2	-2,1
	POSTN	-5,8	-8,0

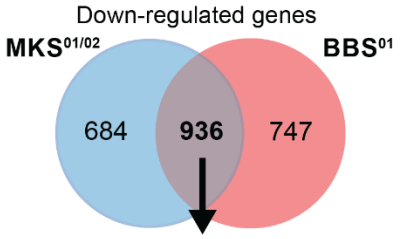
D

Nervous system development (GO:0007399)	Gene Symbol	Log2FC MKS	Log2FC BBS
neurotransmitter loading into synaptic vesicle (GO:0098700)	SLC32A1	6,7	4,1
	SLC17A6	4,2	3,4
	SLC17A7	3,4	3,3
	SLC18A2	2,8	3,2
axonogenesis involved in innervation (GO:0060385)	NTRK1	3,4	4,2
	NPTX1	3,0	2,0
	COL25A1	2,8	2,1
somatic motor neuron differentiation (GO:0021523)	ISL2	5,3	5,3
	ISL1	5,1	4,6
	SOX1	4,6	3,1
	OLIG2	4,1	5,5
regulation of axon diameter (GO:0031133)	NEFL	4,0	2,6
	CNTN2	3,7	4,1
	FOXN4	3,6	4,4

S. Figure 2

Barabino at al. 2020

A



B

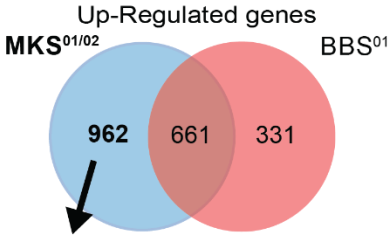
GO Biological process

Cilia and flagella associated (HUGO gene group)	Gene Symbol	Log2FC MKS	Log2FC BBS
	BBOF1	-1,8	-2,5
	CCDC65	-0,8	-1,7
	CFAP100	-1,7	-1,6
	CFAP157	-2,1	-2,5
	CFAP161	-2,7	-3,3
	CFAP221	-1,3	-2,5
	CFAP43	-1,4	-1,7
	CFAP44	-0,8	-1,7
	CFAP45	-3,1	-3,4
	CFAP52	-2,3	-3,5
	CFAP54	-0,9	-1,8
	CFAP57	-3,1	-3,0
	CFAP61	-2,1	-1,9
	CFAP65	-2,2	-3,2
	CFAP77	-3,4	-4,3
	CFAP99	-1,9	-1,9
	DRC3	-0,9	-1,0
	DRC7	-2,5	-2,9
	IQCG	-1,0	-2,0
	ITPR2	-1,2	-1,6
	PPP4R4	-1,4	-2,4

C

GO biological process	Gene Symbol	Log2FC MKS	Log2FC BBS
cellular component organization (GO:0016043)	FOXA3	-1,9	-1,8
	CCDC68	-2,1	-1,5
	GOLGB1	-0,9	-0,7
	MET	-2,5	-1,6
	RAB38	-2,4	-1,3
	USH1G	-2,0	-1,4
	SLC9A3R1	-2,4	-2,1
	PIFO	-2,0	-2,4
	APC	-1,4	-2,1
	TSKU	-2,0	-2,0
cilium assembly (GO:0060271)	MCIDAS	-4,6	-5,3
	FOXJ1	-3,1	-2,6
	CCNO	-4,8	-4,6
	GMNC	-3,9	-5,4
	DRC1	-2,9	-4,1
	TEKT4	-4,8	-2,9
	UBXN10	-2,4	-3,3
	DNAAF1	-2,1	-1,7
	CFAP157	-2,1	-2,5
	PIH1D3	-2,3	-4,1
cilium organization (GO:0044782)	MNS1	-2,8	-2,6
	CFAP61	-2,1	-1,9
intraciliary transport involved in cilium assembly (GO:0035735)	IFT22	-1,4	-1,5
	DYNC2H1	-0,5	-0,7
	KIF3B	-1,1	-1,4
	IFT88	-1,2	-1,1

D



E

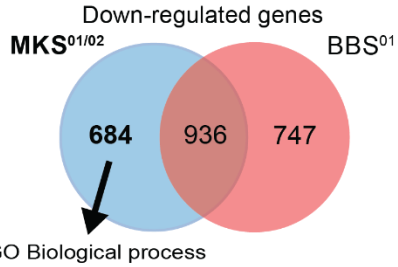
GO Biological process

Developmental process (GO:0032502)	Gene Symbol	Log2FC MKS
anterior/posterior pattern specification (GO:0009952)	GDF3	6,3
	TDGF1	6,1
	HOXB7	5,5
	HOXC4	5,1
	HOXC5	5,0
regionalization (GO:0003002)	PAX8	5,9
embryo development (GO:0009790)	PRDM14	5,1
	APLNR	4,9

F

HOX family	Gene Symbol	Log2FC MKS
	HOXB6	4,3
	HOXB-AS3	4,0
	HOXB5	3,1
	HOXB3	3,1
	HOXA5	3,0
	HOXC6	2,8
	HOXB4	2,7
	HOXB2	2,4

G

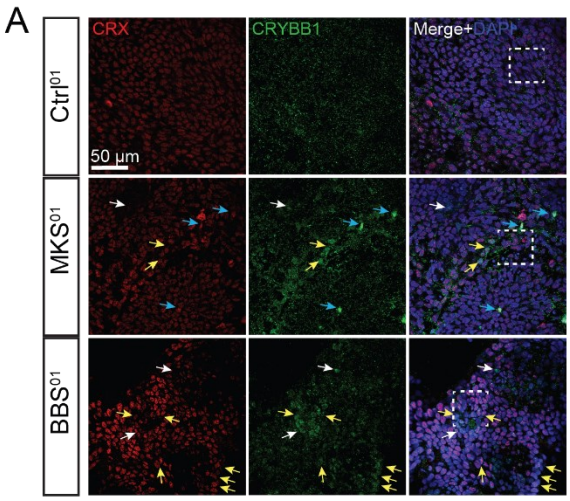


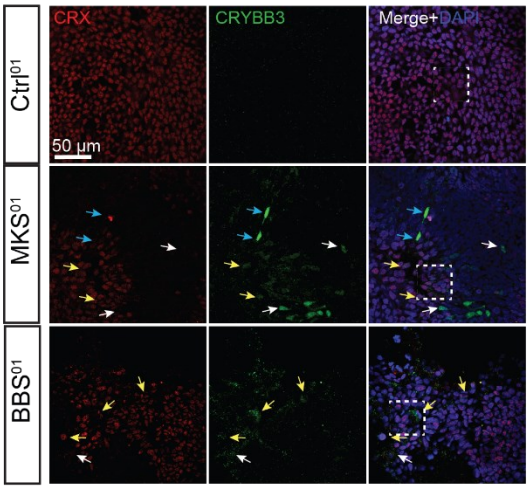
GO Biological process

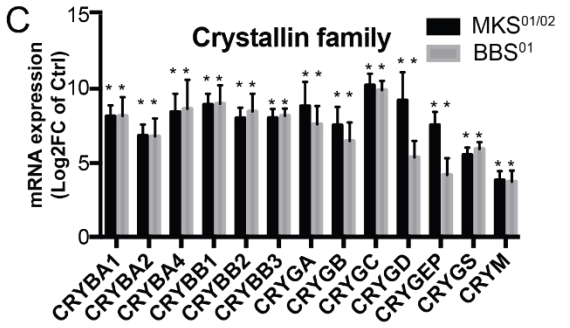
H

Developmental process (GO:0032502)	Gene Symbol	Log2FC MKS
	RPS4Y1	-11,0
	UPK1B	-6,8
	GJB6	-6,4
	GJB1	-5,9
	DHRS9	-5,8
	SPRR1B	-5,5
	PRRX2	-5,4
	NLGN4Y	-5,3
	IVL	-5,2
	CERS3	-5,0
	SPRR1A	-4,7

S. Figure 3

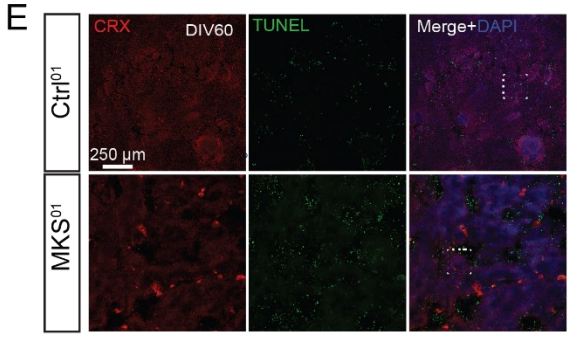


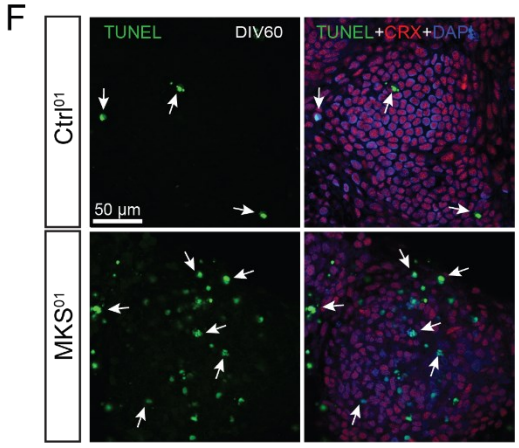


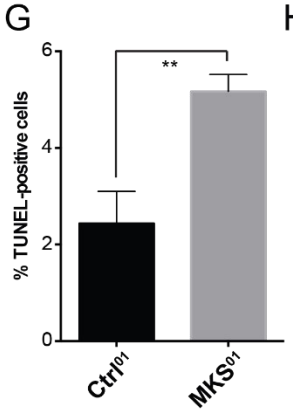


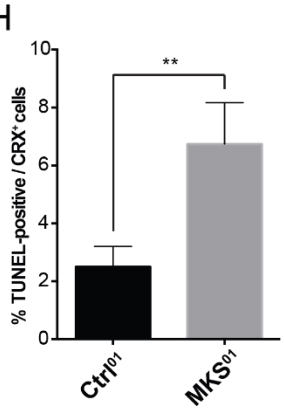
D

Apoptosis	Gene name	Log2FC MKS	Log2FC BBS
apoptotic process(GO:0006915)	ESPL1	2,1	2,7
	DIO3	3,9	2,2
	UNC5A	2,9	3,1
	TNFRSF9	4,9	2,8
	PTH2	7,3	8,7
	FGF8	4,9	3,4
	SFRP2	3,4	4,4
	CDCA7	3,2	2,7
	GADD45G	2,8	3,4
	UNC5C	2,2	2,8
	NSG1	2,8	2,8
positive regulation of neuron apoptotic process (GO:0043525)	PAK3	3,0	0,7
	JUN	2,0	1,4
	CDK5R1	2,4	2,0
	MYBL2	2,3	1,5
	RAPSN	2,3	1,9
	GRIK2	2,4	-0,4
positive regulation of neuron death (GO:1901216)	TLR4	4,8	-1,0
	GRIN2B	2,4	1,2
regulation of neuron apoptotic process (GO:0043523)	AGAP2	3,6	1,8
	MT3	3,0	0,5
	GRIN1	2,4	1,5
	FOXB1	2,0	-1,9
	NRP1	4,2	0,8
	ANGPT1	3,0	-2,8
	DRAXIN	3,2	1,8



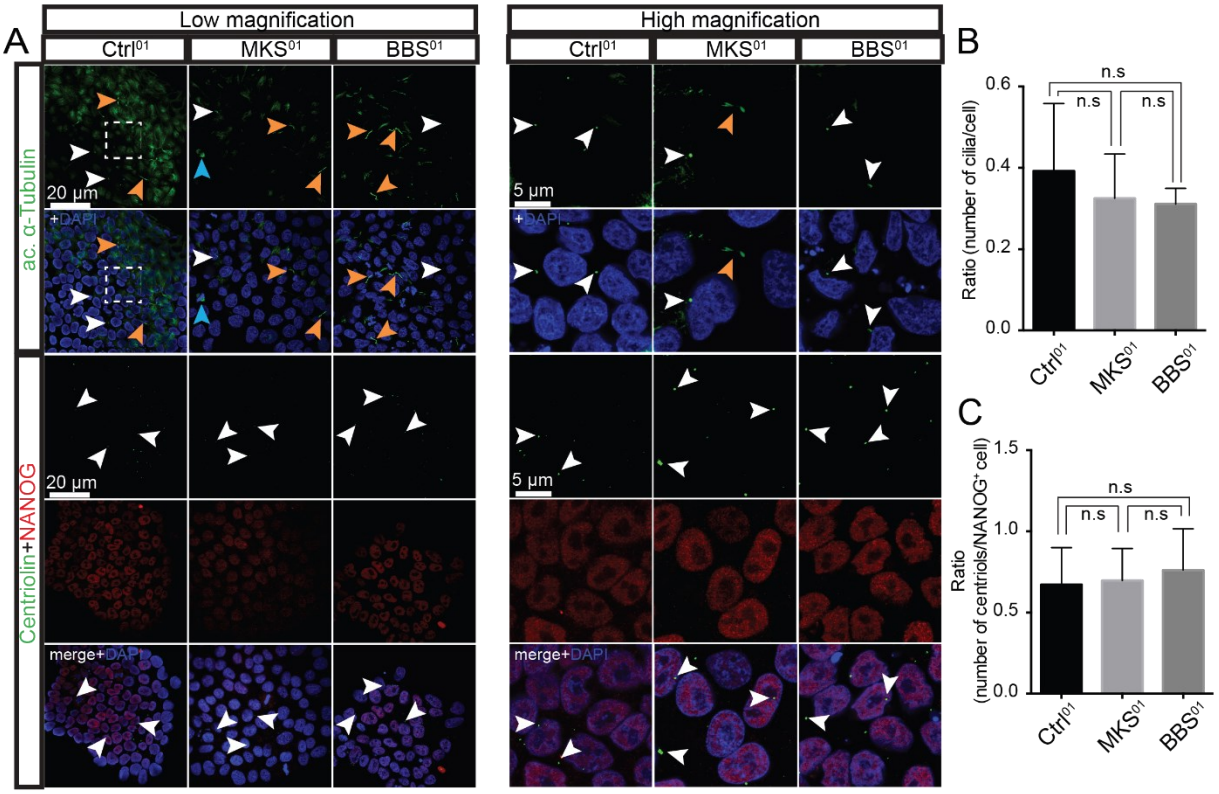






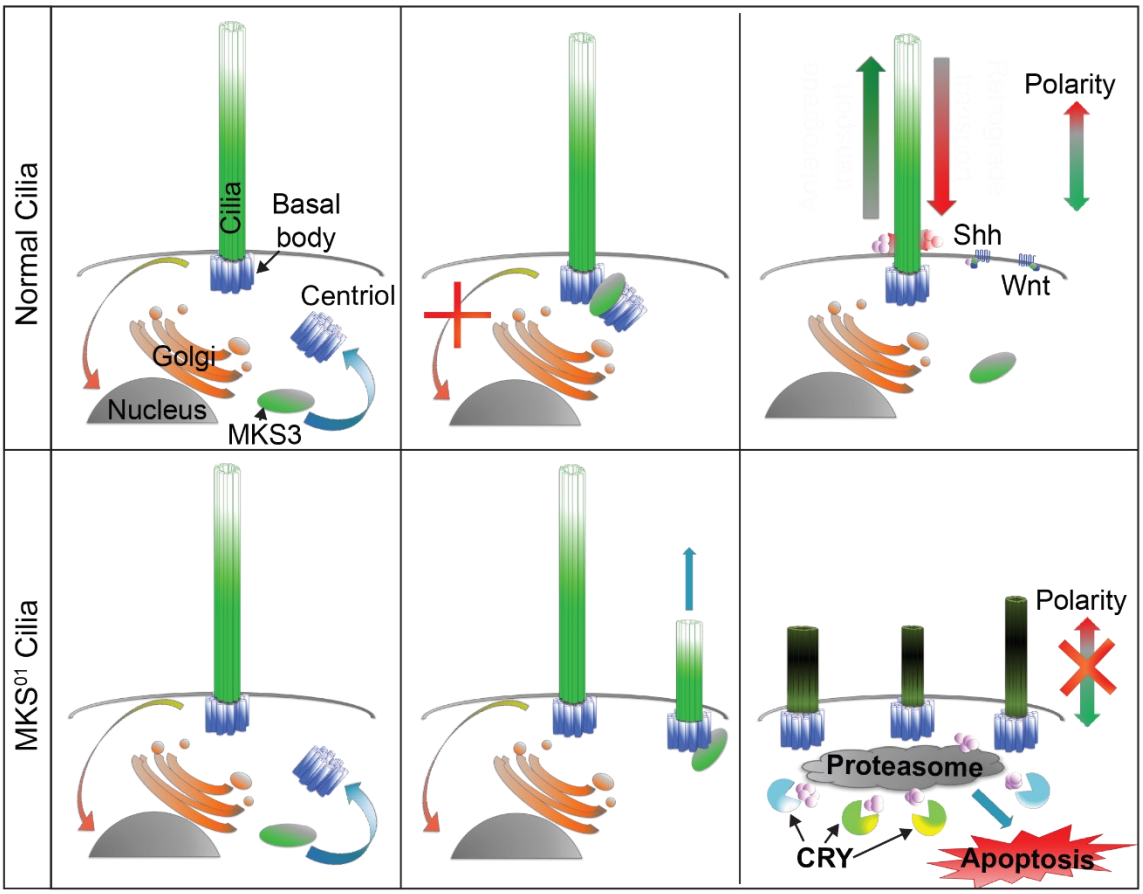
S. Figure 4

Barabino at al. 2020

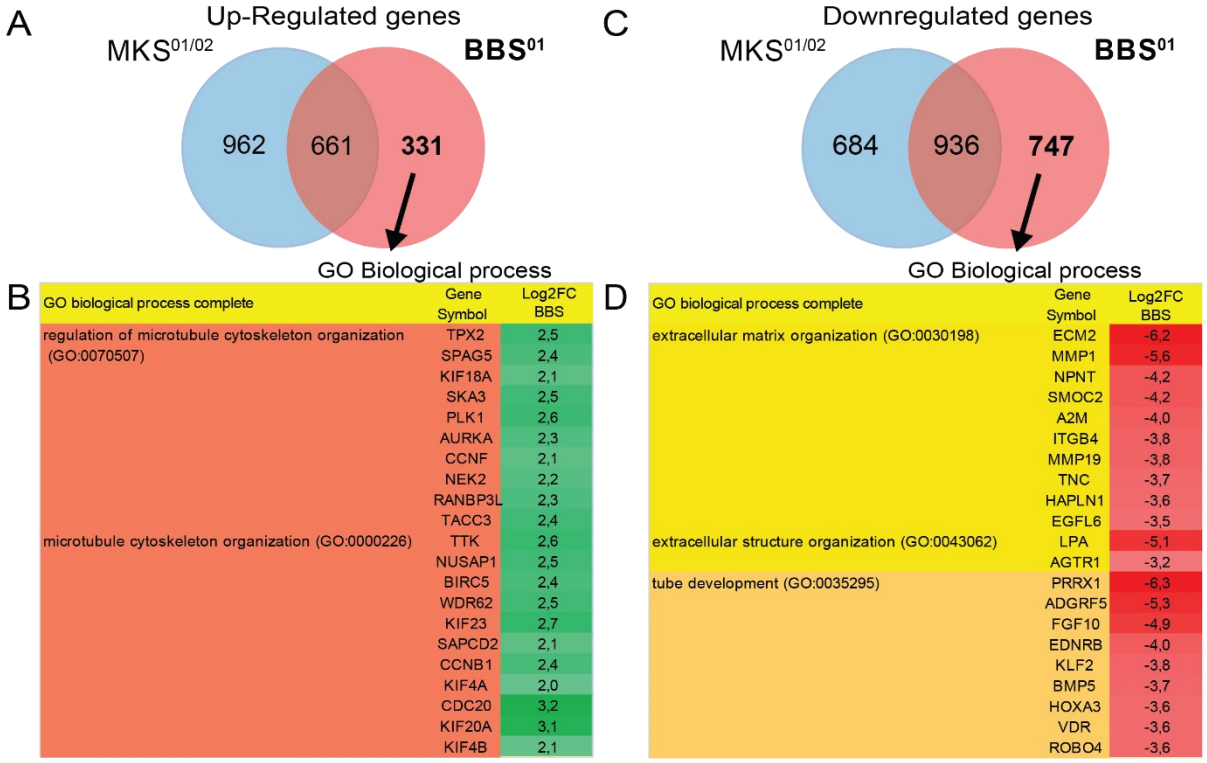


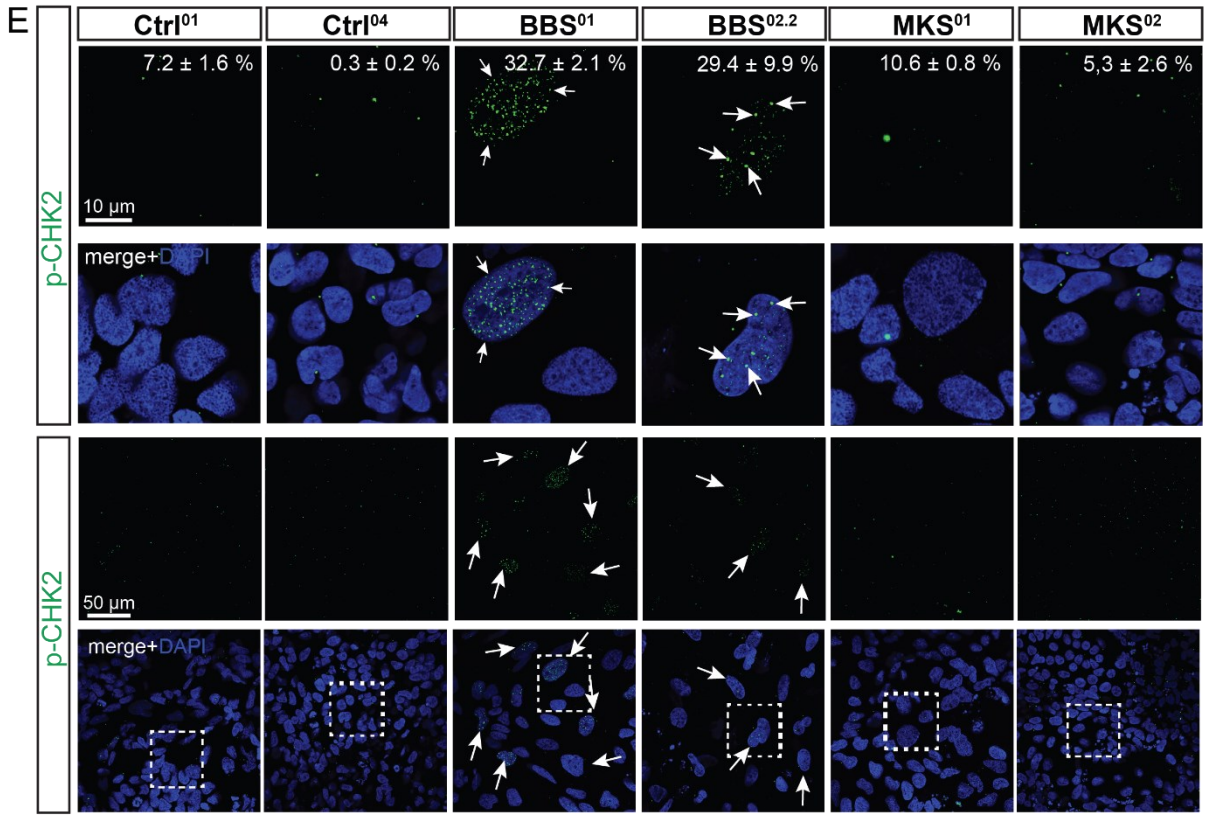
S. Figure 5

Barabino et al 2020



S. Figure 6


A
Up-Regulated genes
MKS01/02
BBS01
962
661
331
C
Downregulated genes
MKS01/02
BBS01
684
936
747
B
GO Biological process
GO biological process complete
Gene Symbol
Log2FC BBS
regulation of microtubule cytoskeleton organization (GO:0070507)
TPX2 2,5
SPAG5 2,4
KIF18A 2,1
SKA3 2,5
PLK1 2,6
AURKA 2,3
CCNF 2,1
NEK2 2,2
RANBP3L 2,3
TACC3 2,4
microtubule cytoskeleton organization (GO:0000226)
TTK 2,6
NUSAP1 2,5
BIRC5 2,4
WDR62 2,5
KIF23 2,7
SAPCD2 2,1
CCNB1 2,4
KIF4A 2,0
CDC20 3,2
KIF20A 3,1
KIF4B 2,1
D
GO Biological process
GO biological process complete
Gene Symbol
Log2FC BBS
extracellular matrix organization (GO:0030198)
ECM2 -6,2
MMP1 -5,6
NPNT -4,2
SMOC2 -4,2
A2M -4,0
ITGB4 -3,8
MMP19 -3,8
TNC -3,7
HAPLN1 -3,6
EGFL6 -3,5
extracellular structure organization (GO:0043062)
LPA -5,1
AGTR1 -3,2
tube development (GO:0035295)
PRRX1 -6,3
ADGRF5 -5,3
FGF10 -4,9
EDNRB -4,0
KLF2 -3,8
BMP5 -3,7
HOXA3 -3,6
VDR -3,6
ROBO4 -3,6


E
Ctrl01
Ctrl04
BBS01
BBS02.2
MKS01
MKS02
7.2 ± 1.6 %
0.3 ± 0.2 %
32.7 ± 2.1 %
29.4 ± 9.9 %
10.6 ± 0.8 %
5,3 ± 2.6 %
p-CHK2
10 µm
merge+DAPI
p-CHK2
50 µm
merge+DAPI

S. Figure 7

SUPPLEMENTARY FIGURE LEGENDS

Supplementary Figure 1. Generation of iPSC cells and identification of the disease-causing mutations from patients affected by BBS and MKS (Related to Figure 1)

(A) Pluripotency marker expression (NANOG, OCT4, and Alkaline Phosphatase Live Stain) in iPSC derived from BBS^{01} and MKS^{01} patients. Nuclei stained with DAPI. (B) Representative images showing the surface of undifferentiated iPSC by scanning electron microscopy (SEM, arrows indicate cilia). (C) Teratoma sections obtained from BBS^{01} and MKS^{01} iPSC. (D) Dissecting microscope images of RPE organoids derived from control and ciliopathy iPS cell lines. (E-G) Schematics of the sequencing results for the four patients ($BBS^{01/02}$ and $MKS^{01/02}$). For each patient are indicated the schematic of the gene affected, the mutated loci with the exon number, the WT sequence in green and the mutated one in red. (H) Summary table of the mutations. For each patient are indicated the patient identifier, the Coriell ID of the fibroblasts of origin, Coriell family, gene name, NCBI Reference number, cChange, and pChange. Scale bars, 100 µm (A; C), 10 µm (B), 2 mm (D).

Supplementary Figure 2. Perturbed expression of Homeobox genes and WNT and NOTCH pathways in MKS and BBS RSs (Related to Figure 2)

(A-D) Gene Ontology Enrichment analyses of common differentially up-regulated and downregulated in the two groups. Some selected significantly enrich GO groups and subgroups are shown with a few representative genes and the respective Log2FC in $MKS^{01/02}$ and BBS^{01} vs $Ctrl^{01/02}$ RSs (P-value <0,05). Log2FC is also represented in a colorimetric scale from red (downregulated) to green (upregulated).

Supplementary Figure 3. Downregulation of genes responsible for cilia assembly, function and maintenance in BBS and MKS RSs and dysregulation of developmental genes in MKS RSs (Related to Figure 3)

(A) Venn diagrams showing the intersection of significant genes differentially downregulated in $MKS^{01/02}$ and BBS^{01} RSs (log2FC>2; P-value <0,05). (B-C) Some representative genes and the respective Log2FC (P-value <0,05) part of the cilia and flagella associated family (B), and other ciliary genes (C). The main GO group and subgroups are shown. Log2FC is also represented in a colorimetric scale from white (0) to red (-5,4). (D-H) Venn diagrams showing the intersection of significant genes differentially up-regulated (D) and down-regulated (G) between $MKS^{01/02}$ and BBS^{01} RSs. (E; F; H) Gene Ontology Enrichment analyses of $MKS^{01/02}$ specific differentially up-regulated (E-F) and down-regulated (H) genes. Selected GO group are shown. Some representative genes and the respective Log2FC (P-value <0,05) are indicated for each group. Log2FC is also represented in a colorimetric scale from white (0) to red (-11).

Supplementary Figure 4. Production of crystallin proteins and activation of apoptosis in MKS and BBS RSs (Related to Figure 4)

(A; B; E) IF representative images used for quantification in figure 4 showing expression of CRX and CRYBB1 (A), CRYBB3 (B), at DIV45 and TUNNEL (E) at DIV60. Blue arrows, cells with strong nuclear expression of Crystallin proteins. Yellow arrows, CRX positive cells showing expression of Crystalline proteins. White arrows CRX negative cells with expression of Crystallin protein. Dotted square, a

representative area used in high-resolution images in figure 4. (C) Gene expression levels of Crystallin family genes significantly upregulated in BBS^{01} and $MKS^{01/02}$ RSs at DIV60. (D) GO group enriched in $MKS^{01/02}$ and BBS^{01} RSs related to the apoptotic process. GO group names and some representative genes and the respective Log2FC. Non-significative Log2FC (P-value>0,05) are in red. (E, F) IF representative images in $Ctrl_{01}$ and MKS_{01} dissociated PRs for CRX and TUNEL. (G, H) Quantification of the percentage of TUNEL-positive cells over total (G) or CRX+ (H) nuclei in $Ctrl^{01}$ and MKS^{01} cells at DIV60 (n = 4 fields for condition, an average of 1057 cells for field). Scale bars: 50 μm (A, B, F), 250 μm (E). All values are means ± SEM. *p < 0.001, **p < 0.01, by Student's unpaired t-test.

Supplementary Figure 5. Expression of Acetyl-α Tubulin and Centriolin in MKS and BBS undifferentiated iPSC (Related to Figure 5)

(A) Expression of Acetyl-α Tubulin, CNTRL, and pluripotency marker NANOG in $Ctrl^{01}$, MKS^{01} and BBS^{01} undifferentiated iPSC. (B-C) Quantification of the number of cilia (Acetyl α-Tubulin) per cells (B) and the number of centrioles (CNTRL) over NANOG positive cells per field (C) in $Ctrl^{01}$, MKS^{01} and BBS^{01} iPSC (n = 5 fields for a condition, an average of 12 cells for a field). Scale bars: 20 μm (A), 5 μm (A'). White arrowhead indicates cilia and centrioles. Orange arrows indicated midbody during cytokinesis that is not considered in the cilia quantifications. All values are means ± SEM., n.s. not significant, by Student's unpaired t-test

Supplementary Figure 6. Mutation in MKS3 protein causes the formation of small supernumerary cilia and aggregation of ciliary proteins (Related to Figures 5 and 6)

Graphic representation of the proposed model for the formation of supernumerary cilia in MKS^{01} photoreceptors.

Supplementary Figure 7. Activation of the mitotic checkpoint proteins and accumulation of DNA damage in BBS photoreceptor progenitors (Related to Figure 7)

(A; C) Venn diagrams showing the intersection of significant genes differentially up-regulated (A) and down-regulated (C) between $MKS^{01/02}$ and BBS^{01} RSs. (B; D) Gene Ontology Enrichment analyses of BBS^{01} specific differentially up-regulated (B) and down-regulated (D) genes. Selected GO group are shown. Some representative genes and the respective Log2FC (P-value <0,05) are indicated for each group. Log2FC is also represented in a colorimetric scale from white (0) to red (-6,3). (E) Expression of pCHK2 in $Ctrl^{01/04}$, $MKS^{01/02}$ and $BBS^{01/02.2}$ photoreceptor progenitors at DIV15. White arrows indicate positive cells. The percentage of positive cells (more than 3 foci per cell) for pCHK2 are indicated on the respective images (average n=159 cells for condition) Dotted squares represent representative area shown in upper images. Scale bar: 10 µm (Top); 50 µm (Bottom).

SUPPLEMENTARY MATERIAL AND METHODS

Generation, Maintenance, and Characterization of iPS Cells

Human PSc were used in accordance with the Canadian Institute Health Research (CIHR) guidelines and approved by the "Comité de Surveillance de la Recherche sur les Cellules Souches" (CSRCS) of the CIHR and Maisonneuve-Rosemont Hospital Ethic Committee**.** Fibroblasts MKS1, MKS2, BBS1, and BBS2 were obtained from clinically diagnosed individuals (Coriell Biorepository) (details of these cell lines and other cell lines used in this paper are in supplementary table) were reprogrammed by two rounds of overnight retroviral transduction with the four Yamanaka factors, *OCT4, SOX2, KLF4,* and *c-MYC*, using the pMIG vector set. Multiple clones from each patient were karyotyped (Cell Line Genetics) and frozen. A subset of these colonies was used for the study. Lines were established on mouse embryonic fibroblast feeders (Global Stem Cell GSC-6001G) in 20% Knock Out Serum Replacement, 1× nonessential amino acids, 1× Glutamax, 1× penicillin-streptomycin, 55 μM β-mercaptoethanol, DMEM-F12, and 10 ng/ml human fibroblast growth factor 2 (FGF2; Invitrogen) and restarted every ~30 passages. For expansion ESC and iPSC were cultured on a Matrigel-coated plate (Corning, Cat# 354277) with a daily change of StemFlex medium according to the manufacturer's instruction (Gibco # A3349401). To generate teratomas, ~ 3 x 10^6 undifferentiated iPS cells were resuspended in 150 μl of a cold 50/50 mix of DMEM-F12 and Matrigel and implanted beneath the neck scruff of NOD-SCID immunodeficient mice. Tumors were harvested 12 weeks later and processed for histology analysis.

Differentiation of human ESC and iPSC into retinal sheets

The differentiation protocol was based on a previous study (Zhou et al., 2015). However, the recombinant COCO was used at 60ng/ml. iPSC and ES were dissociated using ReLeSR™ (Stem cell technologies, Cat# 05872) and platted on growth factor reduced Matrigel (Corning #356231) in StemFlex cell media (Gibco # A3349401) supplemented with ROCK inhibitor (Y-27632;10μM, Cayman Chemical #10005583). Upon 100% of confluency, the media was changed to CI60 Media [DMEM-F12 medium (Invitrogen) containing,1% N2, 2% B27, 10ng/ml IGF1 (peprotech, Cat#100-11), and 5 ng/ml bFGF (peprotech, Cat#AF-100-18B), 60 ng/ml COCO (R&D System, Cat#3047-CC-050), Heparin (Sigma, Cat#H3149]. The medium was changed every day. For single-cell IF analyses, RSs were dissociated for 30min using Cell Dissociation Buffer (Gibco, Cat# 13151014) and resuspended in CI60 media supplemented with ROCK inhibitor (Y-27632; 10μM) and platted on growth factor reduced matrigel coated plates or chamber slides (Ibidi, Cat# 80826).

Fluorescence and Electron Microscopy

Cells were fixed with 4% PFA for 15 min and permeabilized with Triton X-100 for 10 min. Unspecific antigen blocking was performed using 2% BSA in PBST for 30 min. Cells were incubated with the primary antibody overnight at 4C in a humidified chamber (see dilutions in Table). Secondary antibodies were incubated 1h at room temperature. All secondary antibodies were tested alone or in combination to assay for possible non-specific background fluorescence. After incubation with the secondary antibody, slides were counterstained with DAPI. Pictures were taken using a confocal microscopy system (Olympus). Confocal microscopy analyses were performed using 60x objectives with an IX81 confocal microscope (Olympus, Richmond Hill, Canada), and images were obtained with Fluoview software version

3.1 (Olympus). 3D reconstructions were obtained with IMARIS station v8.4.1 (Bitplane). For electron microscopy, structures were scraped from the plate after 5 min of fixation, pelleted at 300g for 4 min, and the pellet was gently released by pipetting into 0.15 M sodium cacodylate trihydrate (Sigma) dissolved in water (pH 7.3) containing 4% formaldehyde and 2% glutaraldehyde (Electron Microscopy Sciences), post-fixed with osmium tetroxide solution (Sigma), dehydrated in serial ethanol dilutions (Sigma) and embedded in epoxy resin. Ultrathin sections (75 nm) were cut, mounted on 200 mesh copper grids, counterstained with uranyl acetate and lead citrate stains (Electron Microscopy Sciences), and examined in a JEOL JEM-1010 transmission electron microscope.

Quantification of immunofluorescence images

High magnification z-stack images of MKS3, Centriolin and acetylated α-Tubulin was used for 3D reconstruction. Using IMARIS station v8.4.1 (Bitplane) software with the Surpass optional module, volume and surface were rendered and quantified (Figure 6D-E). For Figure 5F the channel correlation in colocalized volume option was used to build the colocalization channel and then used for the Coloc Statistics.

RNA sequencing analyses

Total RNA from two independent biological samples (two distinct differentiation) for each cell lines was extracted using the RNeasy Mini Kit (Qiagen, Cat# 74104) according to manufacturer instructions including DNase treatment. RNA integrity was controlled using Bioanalyzer (Agilent) and two samples from each group with a RIN>8 were sequenced. 1 μg of total RNA was subjected to mRNA isolation using OligodT Dynabeads and libraries were prepared using Ion Total RNA-Seq Kit v2. Libraries

were sequenced onto P1 chips from Ion torrent as unpaired to reach 40 million of reads for each sample. Raw sequencing files (FASTQ) were first trimmed to remove adapters and further validated using FASTQC v0.11.7. Reads were aligned onto hg19 using Hisat2 v2.1.0 with default parameters and sorted using samtools v1.9. Gene level counts were determined using featureCounts from subread package. Differential expression analysis was performed using DEseq2 package in R and volcano plot were generated using Tmisc and Calibrate R packages. FPKM values for each sample were calculated using Cufflinks v2.2.1. Heatmaps were generated with FPKM values using Heatmapper2. Gene Ontology was performed using PANTHER (Mi et al., 2010, 2012). Z-scores were computed for each row and clustered using average linkage and Pearson distance measurement method. For Gene-Set Enrichment Analysis (GSEA), differentially expressed genes were pre-ranked according to the log2 of fold change and a p-value < 0.05. Ranked lists were then subjected to GSEA with a classic scoring scheme, an enrichment score normalization and 1000 permutations. Publicly available RNA-seq used in this paper are found in the supplementary table.

Western Blot

Cell extracts were homogenized in the Complete Mini Protease inhibitor cocktail solution (Roche Diagnostics), followed by sonication. Protein material was quantified using the Bradford reagent. Proteins were resolved in 1x Laemelli reducing buffer by SDS-PAGE electrophoresis and transferred to a Nitrocellulose blotting membrane (Bio-Rad). Subsequently, membranes were blocked for 1h in 5% non-fat milk-1X TBS solution and incubated overnight with primary antibodies. Membranes were then washed 3 times in 1X TBS; 0.05% Tween solution and incubated for 1h with corresponding horseradish peroxidase-conjugated secondary antibodies. Membranes

were developed using the Immobilon Western (Millipore). Antibodies used are listed in supplemental experimental procedures.

Whole-genome alignment and variant calling.

Base calling was performed using Illumina HiSeq Analysis Software (HAS; version 2-2.5.55.1311). Reads were mapped to the b37 reference sequence using bwa-mem v0.7.12. Duplicate reads were removed using MarkDuplicates from Picard v2.5.0. Local read realignment around indels, base quality score recalibration (BQSR), variant calling with HaplotypeCaller, and variant quality score recalibration (VQSR), was accomplished using GATK v3.7.0. Resulting variant calls were annotated using a custom pipeline developed at TCAG based on ANNOVAR. Strategy for identification for pathogenic variant: 1. Filter out variants with read depth < 10; 2. Keep homozygous variant in more than 80% of the reads and heterozygous variant with 30%-70% of the reads with variants; 3. Keep variants with =< 1% population frequency; 4. Remove non-coding variants; 5. For coding variants: Keep Synonymous variants that can potentially affect splicing (spx_dpsi =<-4); Keep intronic variants that can potentially affect splicing (spx_dpsi =<-4); Keep loss of function mutation (frameshift substitution, frameshift insertion, non-frameshift substitution, non-frameshift deletion, non-frameshift insertion, stopgain, stoploss); Keep pathogenic missense mutation using pathogenicity scores; Sift =<0.05: Polyphen >=0.95; Ma >= 2; PROVEN =<2.5; ploylopPMam >=1; phylopVert100 >=1.5; CADD phred >=15; 6. Identify variants in cilia-related genes; 7. Use gnomAD and 1000 genome population frequency to further eliminate common variants; 8. Binary alignment map (BAM) validation on IGV.

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER	DILUTION WB	DILUTION IF	HOST
Antibodies					
Fluorescein labeled Peanut Agglutinin (PNA)	Vector laboratories	Cat#FL-1071	N/A	1/250	N/A
S-Opsin	SCB	Cat#sc-14363	1/250	1/200	Goat
RPGR	SCB	Cat#sc-14672	N/A	1/100	Goat
acetilated α-Tubulin	SCB	Cat#sc-23950	N/A	1/500	Mouse
Centriolin (C-9)	SCB	Cat#sc-365521	N/A	1/250	Mouse
MKS3 (EPR17187)	abcam	Cat#ab2020662	1/1000	1/500	Rabbit
CRX (CORD2)	abcam	Cat#ab78662	1/1000	1/300	Rabbit
CRYBB1 (A-8)	SCB	Cat#sc-374496	1/500	1/250	Mouse
CRYBB2 (B12)	SCB	Cat#sc-376006	1/500	N\A	Mouse
CRYBB3 (G-3)	SCB	Cat#sc-374374	1/500	1/250	Mouse
γH2Ax	Millipore	Cat#05-636	1/1000	1/250	Mouse
pChk2	CellSignaling	Cat#2661	N/A	1/100	Rabbit

TKK	SCB	Cat#sc-376842	N/A	1/200	Mouse
PCNA	abcam	Cat#ab18197	N/A	1/500	Rabbit
NANOG	abcam	Cat#ab21624	N/A	1/1000	Rabbit
OTC4	abcam	Cat#ab184665	N/A	1/500	Mouse
53BP1	Novus	Cat#nb100-304	N/A	1/500	Rabbit
GAPDH(D-6)	SCB	Cat#sc-166545	1/500	N/A	Mouse
PolyUBk48 (Apu2)	Millipore	Cat#05-1307	1/2000	N/A	Rabbit
MKS3	proteintech	Cat#13975-1-AP	N/A	1/100	Rabbit

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Biological Samples		
Human eyes	Banque de tissus oculaires pour la recherche en vision (Centre de recherche du CHU de Québec-Université Laval)	NA
Chemicals, Peptides, and Recombinant Proteins		
ROCKi (Y-27632)	Cayman chem	Cat#10005583
Recombinant Human COCO Protein	R&D system	Cat#3047-CC-050
Recombinant human IGF-1	peprotec	Cat#100-11

Recombinant human FGF basic	peprotec	Cat#AF-100-18B
heparin	Sigma	Cat#H3149
Critical Commercial Assays		
TUNEL Alexa Fluor™ 488 Imaging Assay	Invitrogen	Cat#C10245
Deposited Data		
Indifferentiated IPS cells used in RNAseq analisys	GSE87273(Cohen-Hadad et al., 2016); GSM3239684(Sharmin et al., 2016); GSE98288(Hall et al., 2017).	SRX2186365; SRR7458274; SRR7458276; SRX2767163
Indifferentiated ESC cells used in RNAseq analisys	GSM1715067(Jang et al., 2016); GSM3195634(Syrett et al., 2018); GSM2734431(Sun et al., 2018).	SRR2069387; SRR7357230; SRR5905125; SRR5905126.
Human Retina used in RNAseq analisys	GSE94437(Li et al., 2014)	GSM2475287.
RNAseq Data of RSs ($Ctrl^{1/2/3}$, $Ctrl^{4}$(ES), $MKS^{1/2}$, BBS^{1}) (This paper)	GSE133247	GSM3903945; GSM3903946; GSM3903947; GSM3903948; GSM3903949; GSM3903950; GSM3903951; GSM3903952; GSM3903953;

		GSM3903954; GSM3903955; GSM3903956; GSM3903957; GSM3903958.
Whole genome Sequencing of MKS[1] and BBS[1] patients (This paper)	BioSample	SAMN13656497; SAMN13656498.
Experimental Models: Cell Lines		
Ctrl[01]_iPSC	In house (https://doi.org/10.1016/j.celrep.2018.04.097)	N/A
Ctrl[02]_iPSC	Lab. C. Beauséjour	N/A
Ctrl[03]_iPSC	Lonza	N/A
Ctrl[04]_iPSC	Coriell	AICS-0016
Ctrl_ESC (hESC line HUES9)	HSCI iPS Cell Core Facility	Cat#NIHhESC-09-0022
MKS[01] (Fibroblast cell line used to make in house iPSC)	Coriell	Cat#GM07214
MKS[02] (Fibroblast cell line used to make in house iPSC)	Coriell	Cat#GM07817
BBS[01] (Fibroblast cell line used to	Coriell	Cat#GM05948

make in house iPSC)		
BBS^{02} (Fibroblast cell line used to make in house iPSC)	Coriell	Cat#GM05950
BBS$^{02.2}$ (clone 2) (Fibroblast cell line used to make in house iPSC)	Coriell	Cat#GM05950
Software and Algorithms		
IMARIS station v8.4.1	Bitplane	N/A