Roles of Two Lysine Acetyltransferase Activators in Neurodevelopmental Disorders

Negar Mousavi Firouzabadi

Division of Experimental Medicine
Department of Medicine
&
Rosalind and Morris Goodman Cancer Institute
McGill University
Montreal, Canada
December 2023

A thesis submitted to McGill University in partial fulfillment of the requirements for the degree of Master of Science

ABSTRACT

Epigenetic regulation is a fundamental mechanism by which gene expression is governed in all eukaryotic organisms. One type of epigenetic modification is lysine acetylation, which occurs at the N-terminal tails of core histones from yeasts to humans. Mammalian lysine acetyltransferase 6A (KAT6A)- or KAT6B-containing multisubunit complexes are enzymatic transcriptional regulators that acetylate histone H3 at lysine 23 *in vivo*. In the complexes, bromodomain- and PHD finger-containing factor 1 (BRPF1) acts as a scaffold and serves as an activator to stimulate the acetyltransferase activities of KAT6A and KAT6B. Importantly, published papers from this and other laboratories have described heterozygous germline *BRPF1* mutations in about two dozens of patients with syndromic intellectual disability. Through collaborating with various clinicians, we have now identified 17 new patients who possess various *BRPF1* mutations. Some of these novel mutations are located at the coding sequence for the PWWP domain of BRPF1, suggesting the clinical importance of this domain. BRPF2 is paralogous BRPF1, so an important question is whether there are germline *BRPF2* mutations in patients with developmental problems. Related to this, we have identified such mutations in 8 patients with neurodevelopmental and other anomalies, indicative of a new neurodevelopmental disorder. To ascertain the pathogenicity, we have engineered all these 25 *BRPF1* or *BRPF2* mutations. During the course of this research project, we have also optimized a site-directed mutagenesis method and achieved a high efficiency of 50-100%. A manuscript on this optimized method is currently under preparation, which is significant when considering its potentially general and wide usage in different projects in this and many other laboratories. In conclusion, this project has identified and analyzed new *BRPF1* and *BRPF2* mutations, thereby strengthening their direct links to neurodevelopmental disorders.

RÉSUMÉ

La régulation épigénétique est un mécanisme fondamental par lequel l'expression des gènes est régie dans tous les organismes eucaryotes. Un type de modification épigénétique est l'acétylation de la lysine, qui se produit au niveau des queues N-terminales des histones centrales, des levures aux humains. Les complexes multisous-unités de lysine acétyltransférase 6A (KAT6A) et KAT6B de mammifères sont des régulateurs transcriptionnels enzymatiques qui acétylent l'histone H3 au niveau de la lysine 23 in vivo. Dans les complexes, le facteur 1 contenant le bromodomaine et le doigt PHD (BRPF1) agit comme un échafaudage et sert d'activateur pour stimuler les activités acétyltransférase de KAT6A et KAT6B. Il est important de noter que des articles publiés par ce laboratoire et par d'autres ont récemment décrit des mutations germinales hétérozygotes de BRPF1 chez environ deux douzaines de patients atteints de déficience intellectuelle syndromique. Grâce à la collaboration avec divers cliniciens, nous avons récemment identifié 17 nouveaux patients possédant diverses mutations de BRPF1. Certaines de ces nouvelles mutations sont localisées au niveau de la séquence codante du domaine PWWP de BRPF1, ce qui suggère l'importance clinique de ce domaine. BRPF2 est un BRPF1 paralogue, une question importante est donc de savoir s'il existe des mutations germinales de BRPF2 chez les patients présentant des problèmes de développement. Dans le même ordre d'idées, nous avons identifié de telles mutations chez 8 patients présentant des anomalies neurodéveloppementales et autres, révélatrices d'un nouveau trouble neurodéveloppemental. Pour vérifier le pouvoir pathogène, nous avons conçu toutes ces 25 mutations BRPF1 ou BRPF2. Au cours de ce projet de recherche, nous avons également optimisé une méthode de mutagenèse dirigée et atteint un rendement élevé de 50 à 100 %. Un manuscrit sur cette méthode optimisée est en cours de préparation, ce qui est important, compte tenu de son utilisation potentiellement générale et large dans différents projets

de ce laboratoire et de nombreux autres. En conclusion, ce projet a identifié et analysé de nouvelles mutations BRPF1 et BRPF2, renforçant ainsi leurs liens avec les troubles neurodéveloppementaux.

ACKNOWLEDGEMENTS

With a major part right within the Pandemic, my graduate study has been a really challenging journey, but with help from many, I have gone through it successfully. Firstly, I would like to express my sincere appreciation to my supervisor, Dr. Xiang-Jiao Yang, for his exceptional mentorship, patience and wisdom throughout this research endeavor. His expertise, encouragement and insightful feedback have been invaluable in refining my ideas, improving the quality of this thesis and broadening my perspective on the subject. I would like to thank him for his unending patience, support, constant encouragement in fostering critical thinking, and embracing challenges. I am profoundly grateful for the trust that he placed in me, which empowered me to explore new areas and overcome obstacles with confidence. I cannot express my gratitude and thank him enough for the positive impact he has had on me.

Secondly, during the course, we collaborated with Dr. Philippe Campeau at University of Montreal and many other clinicians in different countries. Without their support and help, this project would not have been possible. In addition, the research involves mutations from many clinical cases, so without the support of the patients' and their families, the project would not have been possible either.

Thirdly, I would like to extend my gratitude to the members of my thesis committee, Drs. Peter Sigel, Anastasia Nijnik and Jason Tanny, for their valuable contributions through insightful discussions and meaningful suggestions that have significantly aided the progress of my project. In addition, I need to thank an external reviewer for critically reviewing and evaluating this thesis; the reviewer's singhtful and constructive comements have helped make the final version of this thesis to a much better shape. I would also express my gratitude to Ms Elham Ebrahmini and Ms Arezou Razavi forreading and editing this thesis.

Finally, I am truly thankful to my parents and my sister, who have been the pillars of strength and constant source of inspiration, their unyielding love, unwavering support, and boundless sacrifices have been the foundation upon which I built my dreams and aspirations. Dad: your encouragement during moments of doubt and celebration during moments of success have propelled me forward, reminding me of my potential even when I faltered. Mom: your belief in my abilities has been the driving force behind every achievement in my life, including the completion of this thesis. I am profoundly grateful for the countless sacrifices you two have made to provide me with the best; your unwavering faith in my endeavors has instilled in me the determination to persevere and reach for the stars. Last but not least, I must thank my unbelievably supportive partner Parham. You have demonstrated rare and amazing patience over the last two years.

CONTRIBUTIONS OF AUTHORS

Chapter One

N.M. completed the literature review, designed the figures and wrote the chapters. Dr. Yang edited drafts and N.M. finalized the chapters.

Chapters Two and Three

N.M. and Dr. Yang designed experiments, N.M. wrote the manuscript. Dr. Yang edited drafts and N.M. finalized the chapters.

Chapter Four

N.M. wrote the manuscript, Dr. Yang edited drafts and N.M. finalized the chapter.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF ABBREVAITIONS

Acetyl-CoA: Acetyl-coenzyme A
ATP: Adenosine-5'-triphosphate
BAH: Bromo-adjacent homology
BRPF: Bromodomain-containing protein
CHD: Chromodomain-helicase-DNA-binding protein
CNS: Central nervous system
Co-IP: Co-immunoprecipitation
DDR: DNA damage response
DMEM: Dulbecco's modified Eagle's medium
DNA: Deoxyribonucleic acid
DNMT: DNA methyltransferase
eRNA: Enhancer RNA
FBS: Fetal bovine serum
GCN5: General control of amino acid synthesis 5
HA: Human influenza hemagglutinin
HAT: Histone acetyltransferase
HBO1: Histone acetyltransferase binding to ORC1
HCC: Hepatocellular carcinoma
HDAC: Histone deacetylase
HEK 293: Human embryonic kidney 293 cell line
ISWI: Imitation SWI
ING5: Inhibitor of growth family member 5
INO80: INOsitol requiring 80
PTM: post-translational modification
KAT: Lysine acetyltransferase
KDM: Lysine demethylase
KMT: Lysine methyltransferase
lncRNA: Large non-coding RNA
MEAF6: MYST/Esa1-associated factor 6
MOZ: Monocytic Leukemia Zinc Finger Protein
MORF: MOZ-Related Factor
miRNA: MicroRNA
mRNA: Messenger RNA
MYST: Moz, YBF/SAS3, SAS2, TIP60
ncRNA: Non-coding RNA
PBS: Phosphate buffered saline
PCR: Polymerase chain reaction
PHD: Plant homeodomain
piRNA: Piwi-interacting RNA
PTM: Post-Translational Modifications
PWWP: Pro-Trp-Trp-Pro
REST: Restrictive Element 1-Silencing Transcription Factor
RNA: Ribonucleic acid
rRNA: Ribosomal RNA

Sas: Something about silencing
SDS-PAGE: Sodium dodecyl sulfate-polyacrylamide gel electrophoresis
siRNA: Small interfering RNA
sncRNA: Small non-coding RNA
snoRNA: Small nuclear RNA
SWI-SNF: SWItch/sucrose non-fermentable
TBS: Tris-buffered saline
TIP60: HIV Tat interacting protein, 60 kDa
tRNA: Transfer RNA
WES: Whole exome sequencing

Chapter 1 Literature Review

1.1 Epigenetic Regulation and Human Disease

Deoxyribonucleic acid (DNA) serves as the blueprint of life and is the fundamental basis for the diversity of living organisms. In an eukaryotic cell, its genome in the nucleus is packaged into chromatin. Chromatin is a fiber of nucleosomes, and each nucleosome contains a histone octamer, around which DNA wraps [1]. Throughout every cell cycle in an eukaryotic cell, the chromatin fiber goes through a process of de-compaction and re-compaction, which directly influences the control of gene expression [2]. Accurate gene expression is crucial for different cells and tissues in multiceullar organisms to have the proper functions. Therefore, changes in gene expression may result in irretrievable consequences.

Post-translational modifications (PTMs) are covalent modifications resulting in changed properties of histones or other proteins [3]. These PTMs produce modified proteins, e.g., by proteolytic cleavage, and addition of a modifying chemical group to one or multiple amino acids [3]. To date, over 200 types of PTMs have been discovered, including phosphorylation, acetylation, methylation, and glycosylation [4]. PTMs are also classified into reversible and irreversible modifications. For example, proteolytic cleavage is irreversible, whereas phosphorylation is reversible [5]. PTMs are capable of working alone or coordinating with one another to regulate cellular and systematic processes, resulting in different cell types to arise [3].

1.1.1 Epigenetic Regulation

The term “epigenetics” can be traced back to 1942, when the term was coined by Dr. Conard Waddington, referring to the alterations accruing in the phenotype without changes in the genotype [6]. In a 2008 conference on chromatin-based epigenetics the new meaning of the word

has been established, today being defined as “the study of the stably heritable phenotypes originating from changes in the chromosome, while the DNA sequence has been intact” [7]. In addition, three categories of signals, Epigenators, Initiators, and Maintainers, involved in establishing and sustaining epigenetic states were discussed, emphasizing the importance of understanding these mechanisms in current biological research.

There are four types of epigenetic regulations, i.e., DNA methylation and demethylation, non-coding RNA association, chromatin remodeling and histone modification [8]. These four regulation mechanisms make an impressive impact on gene expression by modifying the state of chromatin to better respond to developmental and environmental stimuli instead of altering the underlying DNA sequence [9].

Due to the continually increasing range of identified PTMs, understanding the intricate patterns and networks of PTMs are crucial to better understand the phenotypes associated with each cell-type identity, paving the way to conduct research on cell programming, chromatin inheritance and designing epigenetic therapies to cure disease. The following section describes the four most common types of PTMs in depth.

1.1.1.1 DNA Methylation

DNA methylation, which used to be known as the most common epigenetic mechanism, is the covalent process in which DNA methyltransferases (DNMTs) transfer and attach the methyl group to the C5 position of the cytosine residue, forming 5-methylcytosine [10] and leading to transcriptional repression [11]. Throughout the process of development, the genome's DNA methylation pattern undergoes modifications caused by a dynamic process that includes methylation and demethylation. This leads to the creation of a distinct and stable DNA methylation pattern in differentiated cells, which in turn controls the transcription of genes specific to each

tissue [10]. Although most of the DNA sites affected by methylation repress gene expression directly, other sites have an indirect effect which would require other complexes or enzymes to compact the chromatin [11].

1.1.1.2 Non-coding RNA Association

Over the past decade, the role of non-coding RNAs (ncRNA) in gene regulation was studied in detail as evidence accumulated proving their role in gene expression [12]. Despite the fact that genomes are broadly transcribed, most of the produced RNA will not be translated into functional proteins but rather have regulatory functions controlling gene expression [12]. The two most well-known ncRNAs are ribosomal RNA (rRNA) and transfer RNA (tRNA), which are in charge of protein synthesis from messenger RNA (mRNA) [12]. To better study ncRNAs, they have been classified into two groups based on their length, small ncRNA (sncRNA) and large ncRNA (lncRNA), having a length of less than or more than 200 nucleotides, respectively [13]. The biogenesis, cellular location, and function of each type of ncRNA is unique which have not been considered in the mentioned classification [14]. Subtypes of sncRNA include microRNA (miRNA), small interfering RNA (siRNA), nuclear RNAs (snoRNAs), and Piwi-interacting RNAs (piRNAs) [15]. Single-strand miRNA has a unique function, resulting in post-transcription gene silencing through translation suppression and endonucleic cleavage enhancement [15]. siRNA has a similar suppression activity through mRNA degradation [16]. In contrast, the activity of enhancer RNAs (eRNA), a lncRNA subtype, results in active gene transcription [17]. As a whole, ncRNAs are capable of altering gene expression either by gene silencing or activation, playing a critical role in epigenetic regulation [18].

1.1.1.3 Chromatin Remodeling

Chromatin remodeling is another form of epigenetic regulation, which packages the chromatin into specific regions: promoters, enhancers, and terminator regions [19]. Despite the condensed structure of chromatin regions, various regions of DNA can still be accessed, and chromatin remodelers play a key role through two mechanisms. Remodeling complexes utilize the stored energy in the ATP to alter histone-DNA structure, resulting in alterations in nucleosomes [19]. To date, four primary families of chromatin remodeling complexes have been identified, SWItch/Sucrose Non-Fermentable (SWI/SNF) family, Imitation SWI (ISWI) family, Chromodomain-helicase-DNA-binding protein (CHD) family, and Chromatin-remodeling ATPase (INO80) family [20]. Despite being classified into different groups, all the remodeling families exhibit common characteristics, such as histone modifications recognition, ATPase regulatory domains, and DNA and nucleosomal affinity. They are capable of disrupting the DNA-histones interactions using DNA-dependent ATPase domains [21].

Remodelers are knowns as versatile tools which help with a wide variety of reactions to change chromatin resulting in altered gene expression, such as nucleosome sliding (sliding an octamer across the DNA nucleosome), histone variant exchange (altering the composition of nucleosomal DNA) [22].

1.1.1.4 Histone Modifications

As described earlier, in eukaryotic cells, DNA is packed into a complex nucleoprotein structure called chromatin [23]. Chromatins are made of smaller unites called nucleosomes, which contain around 147 base pairs DNA wrapped around histone octamers [23]. Histone octamer is formed by two pairs of each core histones (H2A, H2B, H3 and H4) [23]. It is crucial for a cell`s proper function to have this histone octamer correctly formed into the complex with the expection

of the histones N-terminal tail [23]. PTMs can happen on histones in the nucleus by adding or removing moieties from histone tails [24]. These types of modifications can result in either reversible activation or suppression of specific DNA regions. Therefore, this would lead to dynamic control of the cell`s gene expression [25]. This dynamic control of gene expression is not limited to the modifications in the nucleus but also involves all the PTMs that affect cytoplasmic proteins, therby affecting signaling patterns [26]. In the following section, four types of histone modifications will be explained in detail.

1.1.1.4.1 Histone Methylation and Demethylation

As mentioned above, these sets of modifications can occur on DNA and proteins [27]. Lysine methyltransferases (KMTs) and lysin demethylases (KDMs) play a vital role in these modifications [28]. Although all core histones display methylations, Histone H3 is the main site of methylation. Notably, both KDMs and KMTs exhibite a high level of specificity for unique lysin residues and are involved in the extension of lysin methylation [28]. Structural analyses have significantly contributed to our comprehension of KDM biology by identifying specific amino acid residues and their interacting partner [28].

The degree of methylation varies between different lysin residues since they can be mono, di, or tri-methylated (me1, me2, me3, respectively) [29]. Despite the degree of methylation, the electronic charge of the amino acid side chain is intact [29]. Effector molecules, also known as ‘reader’ portions, are capable of recognizing these methylated regions based on the amino acid sequence and their methylation condition due to their methyl-lysine-binding motifs. Some examples of methyl-lysine-binding-motifs are PHD, PWWP, BAH, etc. [29].

The activation or repression of gene expression is based on three factors: the methylation pattern, the degree of methylation, and the specific lysine residue which has been methylated [30].

For instance, H3 Lys-4 trimethyl (H3K4me3) mark activation will result in several genes promoter activation, whereas H3 Lys9 tri-methyl (H3K9me3) will result in gene repression. In order for the gene to be actively transcribed again, the methyl groups must be removed with the help of KDM3A and/or KDM3B [30].

1.1.1.4.2 Histone Phosphorylation

Phosphorylation is the second type of histone modifications, which has been possible due to the function of kinases [31]. Protein kinases form the initiating phosphorylation regions, also known as writers. After the recognition of phospho-binding proteins, also known as readers, the desired modification and effect would happen, before being removed by protein phosphatases [31]. These enzyme proteins can have both activator and inhibitor effect and can show their impact by either directly phosphorylating or indirectly through binding to chromatin [31]. To date, over 500 protein kinases have been discovered and each one has a different degree of selectivity for its substrate [31].

One of the most essential histone phosphorylation's which have been studied in detail is the phosphorylation of Histone H2A at Ser139 (γ-H2AX) [32]. Histone H2A is among the four core histones and is essential for protecting the genome integrity by being involved in DNA damage response [33]. Studies show that the phosphorylation of (γ-H2AX) is one of the immediate responses after DNA damage, aiding in the recruitment of DNA damage repair proteins to the specific damaged region [32].

1.1.1.4.3 Histone Acylation

Histone acylation is another type of modifications that occur on the histones and affect the chromatin structure by neutralizing the positive charge of lysine residues [34]. So far, eight types of short chain acylation modifications have been discovered, affecting the ε-amine group of histone

tail residues, butyrylation (Kbu), hydroxibutyrylation (Khib), propionylation (Kpr), malonylation (Kma), glutarylation (Kglu), crotonylation (Kcr), succinylation (Ksucc), β-hydroxybutyrylation (Kbhb) [35].

The acyl group producers in a cell are result of the cells metabolism and through histone acylation, the cell`s chromatin structure will be adapted to the cell`s metabolism [36]. Histone lysine residue acylations are catalyzed by lysine acetyltransferases (KATs), however, histone lysine acylations are differentiated from acetylation based on specific protein domains [35]. In addition, acylations can be removed by histone deacetylases (HDAC) [37].

1.1.1.4.4 Histone Acetylation

Acetylation is known to be the most studied histone modification [38]. Whenever an acetyl group from Acetyl-CoA is released and is attached to the ε-amine group acetylation happens [39]. KATs are responsible for catalyzing this reaction and the reversible reaction is facilitated by HDACs [39]. Same as acylation, acetylation also works through changing the positive charge of histones to neutral, resulting in lessening the interaction between histones and negatively charged DNA [40]. Loosen DNA is more transcriptionally active; therefore, most acetylation modifications happen on gene promoters and enhancers. Contrarily, when HDACs catalyse deacetylations, transcription would be repressed [40].

Since 1995, which yeast histone acetyltransferase 1 (HAT1) was discovered as the first known KAT, over a dozen KATs were discovered [41]. Most of the KATs are classified into three groups: GCN5-related-N-acetyltransferase (General control of amino acid synthesis 5) (GNAT) family, p300/CREB binding protein (CBP) family and the MYST family consisting of Monocytic Leukemia Zinc Finger Protein (MOZ), yeast Ybf2, yeast Sas2, and mammalian TIP60 [41]. Eighteen types of HDAC have also been discovered and based on their sequence analogy to yeast

HDACs are classified into four groups. Rpd3-like proteins are known as the class I (HDAC1-3/8), Had1-like HDACS known as the class II (HDAC4-7/9/10) [42, 43]. The first two classes are known to have an essential lysin deacetylation function [42, 43]. Class III consists of Sir2-like proteins (SIRT1-7) and class IV has only one member which is the HDAC11 [43]. Class I, II, and IV are zinc dependent amidohydrolases whereas, class II enzymes are divided into class IIa and IIb based on their domain position [42].

Epigenetic modifications are a series of complex and dynamic processes which reflect the interplay between the organism and its environment [44]. The proper function of each component in these complexes is crucial for proper gene expression and the loss of the function will results in various disease types which will be explained in detail in the following sections.

1.1.2 Links of Epigenetic Regulation to Different Diseases

As described earlier, proper epigenetic networks regulations in eukaryotic cells are crucial for gene expression, however, under the effect of environmental factors or developmental stimuli, abnormal epigenetic regulation can occur, resulting in even more severe cases of disease, such as developmental disorders or disease such as cancer [9]. The disease mainly happens through two mechanisms [45]. The first mechanism, the disease is due to the pathological gene expression stemming from a mutation or malfunction of the epigenetic factors [45]. The second mechanism involves collaboration between epigenetic factors and various upstream cellular proteins, particularly DNA-binding transcription factors. The upstream signals exert influence, leading to altered gene expression patterns [45].

A **B**

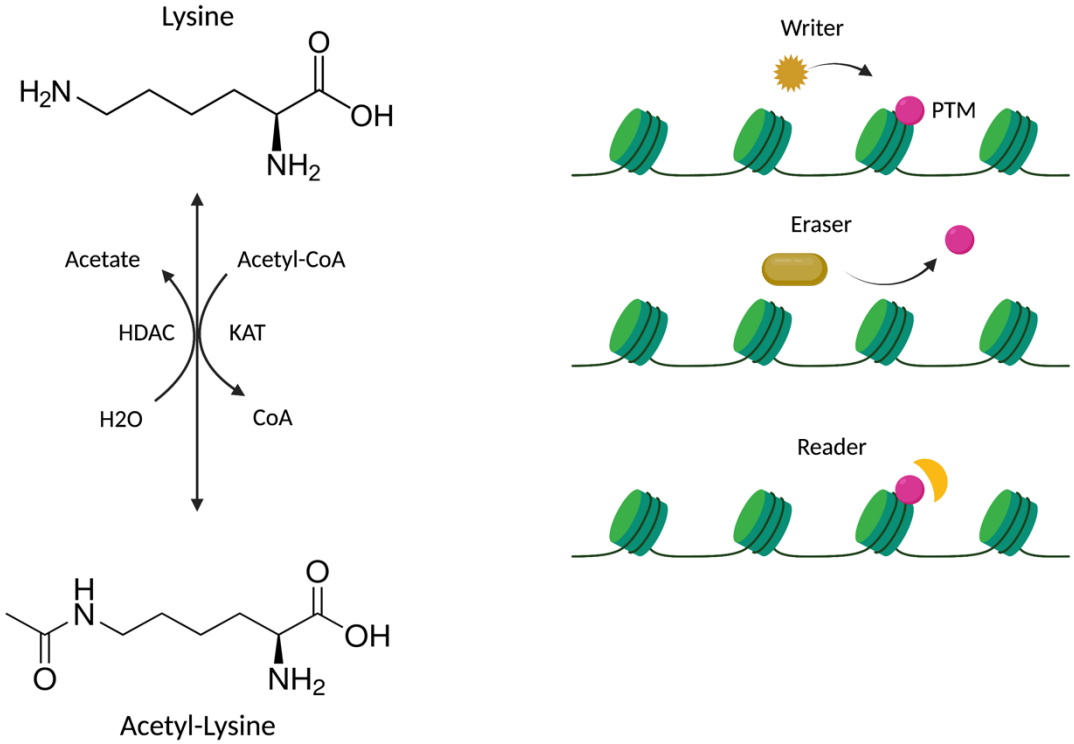


Figure 1.1 Reversible lysine acetylation catalyzed by KATs and HDACs

(A) This panel provides a biochemical depiction of the reversible acetylation process of lysine residues on histones, a crucial post-translational modification that affects chromatin structure and gene expression. The upper schematic represents the structure of a lysine amino acid before acetylation. The enzymes lysine acetyltransferase (KAT) and histone deacetylase (HDAC) facilitate the transfer and removal of an acetyl group (from Acetyl-CoA to the ε-amino group of lysine), respectively. KAT catalyzes the addition of an acetyl group, converting lysine into acetyl-lysine, while HDAC removes the acetyl group, reverting acetyl-lysine back to its original state. The balance between these opposing

enzymatic activities is essential for the regulation of gene transcription, with implications in various cellular functions and disease states.

(B) The second panel illustrates the dynamic interplay of 'writers,' 'erasers,' and 'readers' in the context of chromatin and histone modifications. 'Writers,' such as KATs, are enzymes that introduce post-translational modifications (PTMs) on histones, here depicted as the addition of an acetyl group to a lysine residue. 'Erasers,' like HDACs, remove these modifications, thereby reversing the effects initiated by the writers. 'Readers' are proteins that specifically recognize these PTMs and bind to them, influencing the recruitment of other proteins that regulate transcriptional activity. Together, these interactions dictate the accessibility of chromatin, impacting gene expression patterns and cellular phenotypes. The illustrated representations serve as a simplified visual guide to these complex molecular processes, emphasizing the importance of these enzymatic activities in the maintenance of cellular homeostasis and the potential consequences of their dysregulation. Illustrations created using BioRender.com.

1.1.2.1 Cancer

Cancer, which is among the leading causes of death worldwide, is a result of both dysfunction of genetic and epigenetic regulation [46, 47]. The activation of oncogenes and the silencing of tumor suppressors are the primary shared attributes among all cancer cells [48].In addition, recent studies show that almost every part of cells epigenetic mechanism is affected by cancer and is reprogrammed in favor of the cancerous cells, for instance, DNA methylation, non-coding RNAs, and histone modifications [48].

DNA hypermethylation in CpG-rich promoter regions of tumor suppressor genes is the most popular cancer epigenetic hallmark, however, other regions of the DNA are hypomethylated

[49-51]. Genome-wide hypomethylation leads to oncogene activation and promoting genome instability [51]. Therefore, the chromatin structure will be disformed [51]. As for the histone modifications, studies showed that loss of H4K20me3 and H4K16ac, mediated by HDACs over expression, resulting in loss of acetylation and gene repression [52].

Understanding the epigenetic changes that occur in cancer, has led to the possibility of targeting epigenetic as a therapy method, aiming to reverse the pathological epigenetic functions to normal functions [48, 53]. In recent years, most of the drug discovery was based on targeting DNA methylation and histone modification [48, 53]. Due to the complexity of epigenetics machinery, perhaps the best therapy would be a combinational therapy of two different components of this machinery, however, due to the concerns of potential toxicity of these drugs, further studies are required [48, 53].

1.1.2.2 Neurological Diseases

Dysregulations in cells epigenetic machinery can also lead to neural disorders, specifically alterations in the central nervous system (CNS). Impairments in epigenetic mechanisms have been associated with various neurodegenerative disorders, including Alzheimer's disease [54]. In Alzheimer's disease, there is a dysregulation of the transcription factor known as REST (Restrictive Element 1-Silencing Transcription Factor) [54]. In physiological condition, REST proteins silence stress or apoptotic genes, however in the Alzheimer’s disease, the protein is lost, and the apoptotic genes would be activated as a result [54]. Furthermore,in Alzheimer's disease, there is hypermethylation of specific CpG sites in genes such as ankyrin 1 [54, 55]. Single nucleotide polymorphism (SNP) variants of ankyrin 1 are considered to be risk factors for Alzheimer's disease [55].

1.1.2.3 Developmental Disorders

Numerous clinical studies and research findings have indicated that specific epigenetic elements underwent changes in disorders related to the development of neurons and skeletal structures, primarily in syndromes associated with intellectual disability.

As an illustration, HDAC4 is a type of histone deacetylase that controls crucial genes involved in the development of bones, muscles, the nervous system, and the heart [56]. Haploinsufficiency of HDAC4 has been demonstrated to result in a syndrome called brachydactyly mental retardation syndrome (BDMR) [56]. This syndrome is characterized by intellectual disabilities, developmental delays, and skeletal abnormalities [56].

It is important to highlight that the dysfunction of the MYST family, a group of histone acetyltransferases (HATs) including KAT6A/B, as well as their regulator BRPF1, also leads to various developmental disorders which will be discussed in detail in further sections.

1.2 Histone Acetylation

1.2.1. Overview of Histone Acetylation

Lysine residues can be found at the N-terminal tails of histone proteins and are susceptible to various forms of covalent changes, including acetylation, methylation, and phosphorylation [57]. Histone acetylation, a type of modification, involves the transfer of an acetyl group from acetyl-CoA to the epsilon (ε)-amino group of a specific lysine residue, which is located at an N-terminal tail or other regions of a histone protein. This transfer is catalyzed by an enzyme called histone acetyltransferase (HAT) [58].

The vital role of histone acetylation in activating transcription was first acknowledged in the 1960s [59, 60]. When histones are acetylated, the positively charged epsilon (ε)-amino group on the lysine residue of the core histone tails becomes neutralized [59, 60]. This neutralization

leads to a relaxation of the interaction between the N-terminal tail of histones and the negatively charged phosphate group of the DNA within nucleosomes [59, 60]. As a result, the chromatin structure becomes less condensed, allowing greater accessibility for RNA polymerases and transcription factors [59, 60]. This increased accessibility facilitates the activation of nearby genes and promotes transcription [59, 60]. This is one mechanism whereby histone acetylation affects chromatin structure and gene expression. Another mechanism involves epigenetic readers as discussed below.

Histone acetylation plays a crucial role as it acts as a specific location where epigenetic "readers" can bind to the acetylated histone tails. These "readers" are responsible for stabilizing different components of nuclear signaling machinery and essential processes that rely on chromatin. Proteins containing bromodomains, like histone transferase PCAF/GCN5 and histone acetylation regulator BRPF1, are capable of recognizing acetylated histones. Additionally, PHD fingers can also serve as reading domains for binding to acetyl lysine. Previously obtained results at this lab has demonstrated that the acetylation of histone H3 at either lysine 9 or lysine 14 promotes the binding of MOZ/MORF proteins, which possess tandem PHD fingers, to histone H3 [61]. This binding suggests that the PHD finger-containing MOZ/MORF proteins have the ability to recognize acetyl lysine [61].

1.2.1.1. Classification of Histone Acetyltransferases

In 1999, the initial discovery of an acetyltransferase came from the ciliate protozoan *Tetrahymena thermophila* [62]. Since then, researchers have been able to identify and classify various histone acetyltransferases (HATs) based on their sequence similarities [63]. While HATs are evolutionarily conserved from yeast to humans, only a limited number of human HATs, approximately a dozen, have been identified thus far [64].

Based on sequence similarity, mammalian histone acetyltransferases (HATs) can be categorized into three primary groups [65]. The first group is the GCN5-related N-acetyltransferases (GNATs) family, which is named after its founding member GCN5 [65]. This family includes other members such as PCAF, Elp3, Hat1, Hpa2, and Nut1143 [65]. The second group is the p300/CBP (CREB-binding protein) family, which has been extensively studied and characterized [65]. Lastly, the third group is the MYST family, which is the main focus of this thesis. The name MYST is derived from its four founding members: human MOZ14, yeast Ybf2 (renamed Sas3), yeast Sas2, and human TIP60 [65]. Furthermore, several potential acetyltransferases (such as Spt10) exist that possess motifs resembling those discovered in HATs. However, it is yet to be verified whether these acetyltransferases indeed exhibit acetyltransferase activity [65]. While the MYST and GNAT families differ substantially in their primary sequence, the majority of histone acetyltransferases (HATs) possess a common acetyl-CoA binding domain or Motif A, as well as a structurally similar active site [65].

The human MYST family comprises five members that play a crucial role in maintaining histone acetylation and the overall lysine acetylome [66]. This family consists of two founding members, MOZ (monocytic leukemia zinc finger) and TIP60 (HIV Tat-interacting 60 kDa protein), as well as three newer members: hMOF (homolog of males-absent on the first), HBO1 (HAT bound to ORC1), and MORF (MOZ-related factor) [66]. In the last two decades, MYST family proteins have undergone two systematic renaming processes, such as MOZ (also known as MYST3 and KAT6A), MORF (also known as MYST4 and KAT6B), HBO1 (also known as MYST2 and KAT7), hMOF (also known as MYST1 and KAT8), and TIP60 (also known as KAT5), have undergone alternative naming conventions [67].

1.2.1.2. Molecular Functions of Histone Acetylation

Histone acetylation is crucially involved in controlling chromatin assembly, transcriptional activation, DNA repair, and various other DNA-based processes. This occurs by modifying histone lysine side-chains and other transcription factors through acetylation [63].

One major molecular function of histone acetylation in transcription activation. The initial discovery of the histone acetyltransferase GCN5 revealed its role in acetylating histone H3 at lysine 9 and lysine 14 (H3K9ac and H3K14ac), thereby facilitating transcription [68]. These specific histone acetylations define distinct regions within chromatin that are permissive for gene activation [68]. The level of histone acetylation in living organisms is carefully regulated due to the reversible nature of this modification. Histone deacetylation, which involves the removal of acetyl groups from lysine residues, serves as a counterbalancing process to maintain tight control over histone acetylation levels [69]. It is noteworthy that increased expression of histone acetyltransferases leads to elevated histone acetylation, while mutations [69] of HATs has been associated with various human diseases [69].

The other two important functions of histone acetylation are DNA repair and DNA replication. The initial indication of the involvement of HATs in DNA repair in mammalian cells came from the observation that p300, a HAT, interacts with PCNA (proliferating cell nuclear antigen). PCNA plays a crucial role in both DNA replication and DNA repair processes. This interaction between p300 and PCNA suggests that HATs are engaged in chromatin remodeling and DNA repair synthesis [70]. Other studies also show that HBO1 complex dysfunction also result in reduction of DNA replication [71].

1.2.1.3. Biological Functions of Histone Acetylation

Histone acetylation is increasingly being acknowledged, due to its ability to incorporate environmental signals at the cellular level. As we continue to deepen our understanding of its role

in mouse development and human disease, its significance becomes increasingly apparent. Instances in mice include: the H3K14 acetylation which is essential for fetal liver erythropoiesis and BRPF1 deactivatin result in neural tube closure and vascular defects leading to death of embryo [72, 73]. Some examples of the association between histone acetylation dysregulation and human disease are aggressive forms of leukemia and blood malignancies arising from KAT6A/B dysregulation caused by chromosomal translocations [74]. Neurodegenerative disorders like Alzheimer's Disease (AD) and Parkinson's Disease (PD) have also been observed as a result of histone acetylation dysfunction [75, 76]. The development of epigenetic therapeutics has promise for disease treatment and even utilizing them as biomarkers for diseases such as depression [77]. The urgent concern in therapy and clinical diagnoses revolves around gaining a thorough comprehension of histone acetylation and other epigenetic modifications.

1.2.2. The MYST Family of Human Histone Acetyltransferases

In humans, the MYST family consists of five members, which are MOZ (also known as MYST3 and KAT6A), MORF (also known as MYST4 and KAT6B), HBO1 (also known as MYST2 and KAT7), hMOF (also known as MYST1 and KAT8), and TIP60 (also known as KAT5). While each member contains the enzymatic MYST domain and possesses the ability to acetylate various histone substrates, they differ significantly in their domain organization, formation of complexes, and biological functions [66]. The following section will provide a brief overview of the MYST families, highlighting their molecular characteristics and their roles in mouse development and human diseases.

1.2.2.1. MOZ and MORF (KAT6A and KAT6B)

The MOZ gene was first identified in 1996, as a study linking it to acute myeloid leukemia (AML) and chromosome translocation [78]. The MORF gene was uncovered in 1999 by

conducting a BLAST search on EST databases, utilizing the amino acid sequence of MOZ as a probe [79]. Both are large (~240 kDa) proteins with similar domain organization and characteristics (Fig. 1.2). The proteins MOZ and MORF have the ability to combine with histone reader BRPF1 (Bromodomain- and PHD finger-containing protein 1), along with two additional subunits, to form a tetrameric complex [80]. BRPF1 acts as a scaffold, binds to the MYST domain and significantly enhances the acetylation activity of MOZ [80, 81]. Furthermore, BRPF1 can interact with ING5 (inhibitor of growth 5) and EAF6 (homolog of yeast Esa1-associated factor 6) through its EPC-II motif [73]. The presence of ING5 and EAF6 further amplifies the stimulation and stability of the complex [73]. Hence, BRPF1 fulfills a dual role by functioning as a structural support for complex assembly while also operating as an epigenetic regulator, promoting acetylation, and influencing the specificity of histone substrates [73].

In the MOZ/MORF structures, N-terminal region contains tandem PHD fingers and the MYST domain. Interestingly, Drosophila Enok which plays a critical role in neuroblast proliferation and the development of oocytes in the female germline, N-terminal region bears a striking resemblance to the MOZ/MORFs protein [82, 83]. It has been demonstrated that MOZ and MORF enzymes exhibit their enzymatic functions by acetylating histone H3 at specific lysine residues, namely lysine 9 and lysine 14 [84]. In addition, the tandem PHD fingers of MOZ/MORF play a critical role and are essential for the recognition and binding to acetylated histone H3 at lysine 14. This interaction enables the localization of MOZ/MORF to chromatin [85].

The C-terminal to the MYST domain in MOZ/ MORF consists of ED (glutamate/aspartate-rich) and SM (serine/methionine-rich) domains, both being present in zebrafish orthologs (Fig. 1.2). Although the function of ED domain is still not determined, SM domains is known to be a coactivator for DNA-binding transcription factors [41]. Aberrant MOZ and MORF proteins have

also proven to be the underlying cause of abnormal mouse development and human disease which will be reviewed later in this chapter.

1.2.2.2. HBO1 (KAT7)

Initially detected through a yeast two-hybrid assay as a protein interacting with the DNA replication initiator subunit ORC1, the histone acetyltransferase HBO1 was identified as a member of the MYST family KATs [86]. Despite having a much smaller size compared to MOZ and MORF HBO1, an essential contributor to histone H3 and H4 acetylation, plays pivotal roles in both transcription and the regulation of DNA replication [86]. Research findings have indicated that HBO1 can be found in two distinct forms of multisubunit complexes. These complexes consist of ING4/5, hEaf6, and either scaffold proteins JADE1/2/3 or BRPF1/2/3 [87, 88]. While both types of scaffolds provide HBO1 with the capability to acetylate both H3 and H4 in core histones, they exhibit contrasting specificities in chromatin. The HBO1 complexes containing JADE1/2/3 show a preference for acetylating histone H4, whereas the complexes containing BRPF1/2/3 primarily target histone H3 for acetylation within nucleosomes [89]. Although it has been shown that HBO1 deficiency leads to mice embryonic lethality, there has been no direct link disease found for human HBO1 deficiency [90]. More studies are needed to expand our understanding of the HBO1 and its association with human disease.

1.2.2.3 TIP60 (KAT5)

TIP60, one of the well studied members of the MYST family, was TIP60 was initially discovered as a protein which has specific interaction with the activation domain of the human HIV Tat protein [91]. TIP60 plays a role in numerous cellular processes including transcriptional regulation, apoptosis, autophagy control of the cell cycle, and DNA repair [92]. The primary sites of acetylation on TIP60 include lysine (K) 5, 8, and 12 on histone H4, as well as K5 on histone

H2A and the histone variant H2AX [92]. TIP60 dysfunction has been proven to have direct link to different types of human disease. Recently, a study reported three individuals who exhibited developmental delay, cerebellar malformation, seizures, intellectual disability, sleep disturbance, and epilepsy as a result of heterozygous *de novo* mutations in the TIP60 gene [93]. These findings suggest that TIP60 activity may play a crucial role in neurogenesis and neuronal functions [93].

TIP60 is involved in the acetylation of various non-histone targets with a significant number of these targets being transcription factors [94]. One non-histone substrate of TIP60 is the protein p53, which acts as a tumor suppressor [94]. TIP60 specifically targets lysine 120 of p53, thereby influencing the decision between cell cycle arrest and apoptosis following p53 activation [94].

1.2.2.4 hMOF (KAT8)

MOF, known as hMOF/MYST1/KAT8 was first identified in Drosophila as a component of the dosage-compensation complex, which is responsible for regulating the gene dosage on the X chromosome [95]. Later, it was discovered that hMOF (human MOF) is required for H4 lysine 16 acetylation [96]. Furthermore, biochemical purifications have revealed that KAT8 is associated with two multiprotein complexes: the male-specific lethal (MSL) complex and the KAT8 regulatory nonspecific lethal (KANSL) complex. These complexes are responsible for the acetylation of histone H4 at the lysine 16 position (H4K16Ac). Additionally, the KANSL complex has the ability to acetylate other lysine residues on histone H4, such as H4K5 and H4K8 [97, 98]. hMOF also performs acetylation of p53 at lysine 120, thereby influencing the regulatory mechanisms of p53 in cell cycle arrest and apoptosis [99].

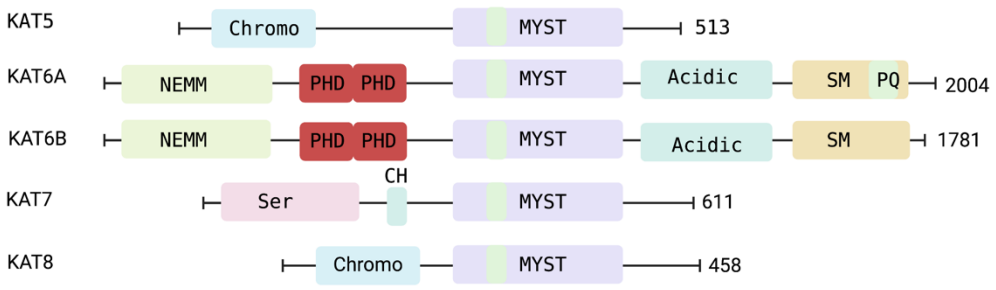


Figure 1.2 Domain Architecture of the MYST Family Histone Acetyltransferases

Schematic representation of the MYST complexes: KAT5, KAT6A, KAT6B, KAT7 and KAT8. This figure provides a schematic representation of the domain organization of various members of the MYST family of histone acetyltransferases (HATs), specifically KAT5, KAT6A, KAT6B, KAT7, and KAT8. Each enzyme is depicted with distinctive functional domains: chromodomains with the potential for binding to methylated histones, NEMM domains, plant homeodomain (PHD) fingers implicated in the recognition of specific histone tail modifications, serine-rich (Ser) regions, MYST domains characteristic of the MYST family with acetyltransferase activity, and acidic regions. The numbers at the ends of each schematic denote the total amino acids in each protein. The visualization clarifies the common features shared by the MYST HATs, as well as the unique elements specific to each KAT, reflecting the specificity in histone acetylation and roles in the regulation of gene expression. The structural diversity illustrated by the presence of additional domains and varying lengths contributes to the functional versatility and differential regulatory mechanisms employed by each HAT within the MYST family. Image generated with BioRender.com.

1.2.3. PZP and EPC module-containing histone readers in MYST complexes

Numerous histone writers have the ability to form functional complexes with multiple reading domains. This allows them to carry out histone modifications. A prominent example of

this can be seen in the complexes containing histone acetyltransferases MOZ and MORF168. In these complexes, the scaffold subunit BRPF1 possesses multiple reading domains, including PHD fingers, bromodomain, and PWWP domain [61]. In MYST family complexes, the epigenetic regulators involved in histone and DNA binding consist of proteins from the BRPF1/2/3 (BRPFs) family and the JADE1/2/3 family. These proteins possess a highly homologous PZP (PHD-zinc knuckle-PHD) domain [67]. Additionally, their EPC-I and EPC-II domains exhibit significant similarity to the EPC1 protein found in the TIP60 complex. More details about these complexes will be reviewed below.

1.2.3.1. BRPF1 and its homologs

In 1994, a protein known as BR140 (bromodomain protein with an estimated molecular weight of 140 kDa) was identified and cloned, later recognized as BRPF1 (Bromodomain- and PHD finger-containing protein 1) [100]. Through sequencing analysis, it was discovered that BRPF1 functions as a zinc finger protein that possesses a bromodomain. The bromodomain is widely recognized as the most extensively studied reader protein for acetyl-lysine interactions [101]. BRPF1 is a versatile epigenetic regulator that encompasses several structural modules, including three modules for binding to histones and one module for binding to non-specific DNA [61]. The first PHD finger in the N-terminal PZP module of BRPF1 specifically recognizes the unmodified N-terminus of histone H3. The C2HC knuckle and the second PHD finger function as a module that binds to the DNA backbone. The central bromodomain of BRPF1 has the ability to bind to acetylated lysine residues. Additionally, the C-terminal PWWP domain can bind to methylated histone H3 [61]. BRPF1 also contains two EPC (Enhancer of Polycomb)-like motifs that facilitate the formation of acetyltransferase complexes [61]. BRPF1 has been demonstrated to interact with and activate three members of the MYST family, namely MOZ, MORF, and HBO1.

The first EPC-like (EPC-I) motif of BRPF1 interacts with these proteins through a conserved N-terminal region. On the other hand, the second EPC-like (EPC-II) motif of BRPF1 functions as a scaffold within the complex, facilitating its interaction with two small non-catalytic subunits, ING5 and hEAF6. This interaction contributes to the stabilization of the tetrameric complex [61]. BRPF2 and BRPF3 are BRPF1s paralog, however, in contrast to BRPF1, both BRPF2 and BRPF3 exhibit a preference for forming complexes with HBO1 rather than MOZ and MORF, despite their significant sequence similarity to BRPF1 [67].

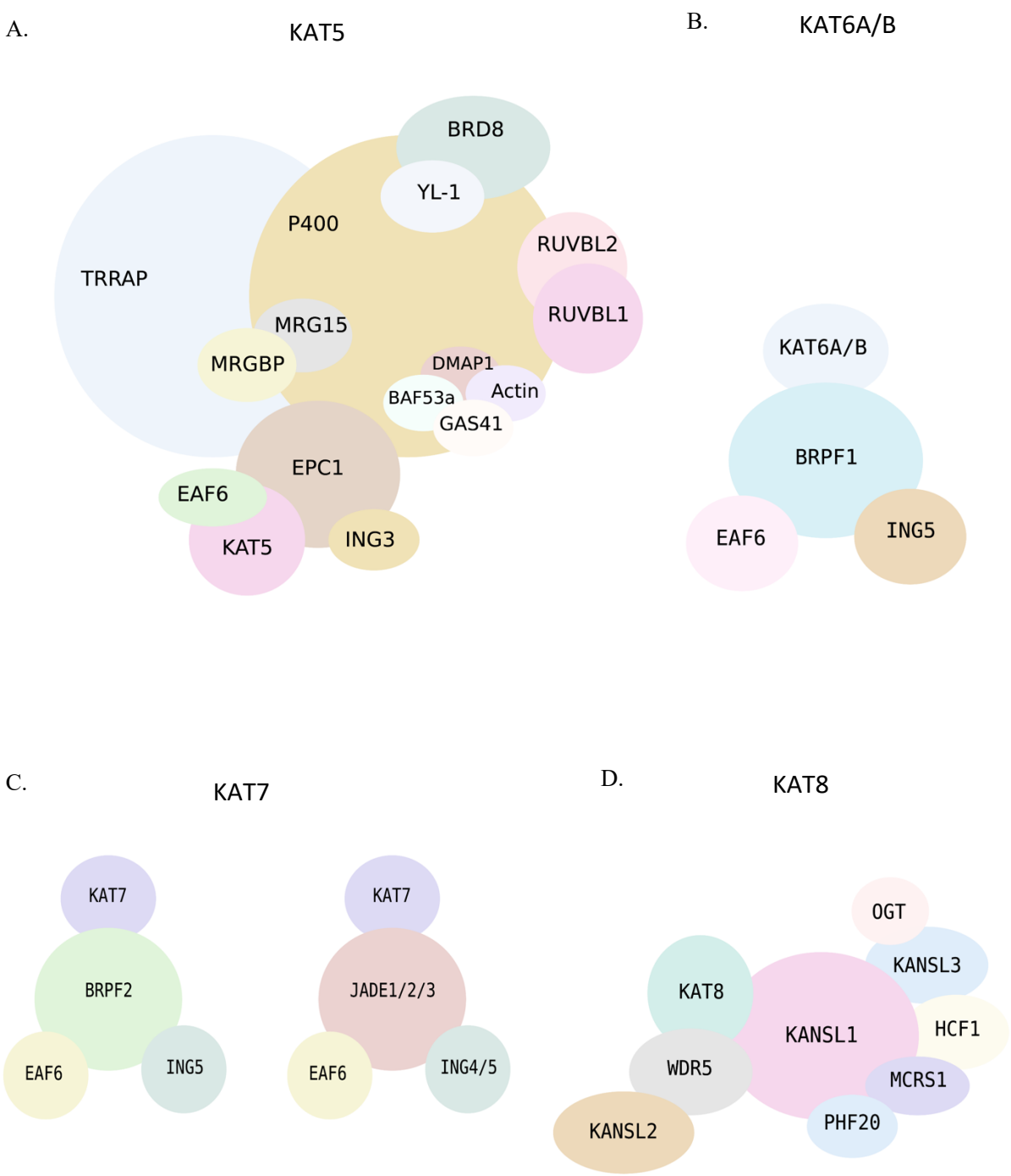


Figure 1.3 Subunit Composition of the MYST Protein Complexes

Schematic representation of the MYST complexes: KAT5, KAT6A, KAT6B, KAT7 and KAT8. This diagram visually dissects the subunit composition and interaction partners for the MYST family of protein complexes, including KAT5, KAT6A/B, KAT7, and KAT8. Panels A to D highlight the individual members of the MYST family and their associated proteins, indicating the complexity of interactions within cellular pathways. Notably, the figure underscores the multifaceted nature of these complexes, with each colored oval representing a distinct set of protein-protein interactions that contribute to the overall function of the MYST acetyltransferases Image generated with BioRender.com.

The PWWP domain is a protein module identified in eukaryotic nuclear proteins that typically appears alongside domains involved in the establishment or recognition of histone modifications. It is characterized by a conserved Pro-Trp-Trp-Pro motif [102]. The focus of my first project is on PWWP domain of BRPF1 as there has not been previous studies investigation the effect of mutations on the PWWP domain. Mutations in proteins containing the PWWP domain have been associated with various human diseases. For example, the WHSC1 gene is situated within the Wolf-Hirschhorn syndrome critical region on chromosome 4p16.3 and is disrupted by chromosomal translocation in cases of lymphoid multiple myeloma disease [103]. Recent studies have demonstrated an association between mutation in BRPF2 PWWP domain and both schizophrenia and bipolar affective disorder [104].

1.2.3.2. JADE1 and its homologs

While the bromodomain is absent, the conserved PZP module and two EPC-like motifs are present in JADEs (proteins encoded by genes associated with apoptosis and epithelial differentiation). JADE1, the initial member of the family, was initially discovered as a protein that

interacts with the von Hippel-Lindau tumor suppressor to regulate the cellular sensing of oxygen levels [105]. It is now determined that JADE1 plays a crucial role in the assembly of the HBO1-HAT complex, which is responsible for histone H4 acetylation. The N-terminal region of JADE1, referred to as region I, is responsible for recruiting HBO1, while the C-terminal region of JADE1, known as region II, is responsible for recruiting the ING-MEAF6 subcomplex. This coordinated recruitment enables the formation of the tetrameric HBO1-HAT complex and facilitates histone H4 acetylation [87]. Both JADEs and BRPFs can form tetrameric complexes with HBO1, ING4/5, and EAF6 [72].

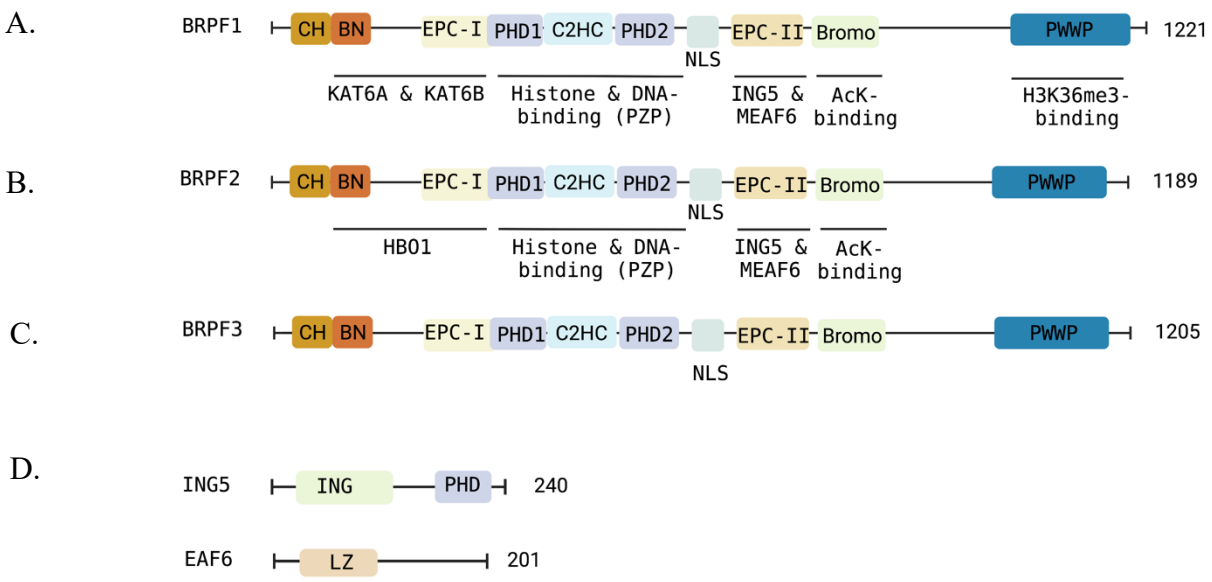


Figure 1. 4 Domain organization of BRPF1 and its paralogs

Schematic representation of BRPF1/2/3, ING5, and EAF6. This figure provides a detailed schematic of the domain organization of the three BRPF (Bromodomain and PHD finger-containing) proteins, BRPF1, BRPF2, and BRPF3, along with ING5 and EAF6, highlighting their modular structures and potential functional domains. For BRPF1 (Panel A), BRPF2 (Panel B), and

BRPF3 (Panel C), key domains such as bromodomain, PWWP, and PHD fingers are illustrated, highlighting their roles in DNA/histone binding and their involvement in the regulation of chromatin structure. Notably, the presence of specialized domains such as the bromodomain suggests an epigenetic regulatory function, interacting with acetylated lysine residues. ING5 (Panel D) possesses ING and PHD domains, the latter of which is implicated in recognizing histone marks and critical for the regulation of gene expression, while EAF6 features a leucine zipper (LZ) motif, which may mediate protein-protein interactions. Each protein is illustrated with its respective amino acid length, providing a relative comparison of their sizes. The unique domain architectures suggest diverse roles in epigenetic regulation, with specific emphasis on the recognition and binding of histone modifications, DNA, and other proteins involved in chromatin remodeling. Image generated with BioRender.com.

1.2.4. KAT6 or BRPF1-related diseases

The main biological role of MOZ (KAT6A) was initially discovered through its involvement in a chromosomal translocation in acute myeloid leukemia (AML) [73]. The MORF gene is similarly involved in fusion with the CBP gene, and this fusion has been associated with acute myeloid leukemia [106]. Additionally, MORF is frequently mutated in leiomyoma, prostate cancer, and breast cancer [107]. MOZ and MORF mutations will not only lead to hematological malignancies, but also solid tumors such as breast cancer [108]. In addition, MOZ and MORF both showed that they play an essential role in various developmental stages such as hematopoiesis [109]. Mutations in MORF cause different types of developmental disorders and mutations in MOZ can also cause syndromes and they all share the common feature of intellectual disability [110-114].

BRPF1 is evolutionarily conserved across different species, ranging from Caenorhabditis elegans to humans [115]. Dysfunction of BRPF1 in zebrafish leads to changes in anterior Hox gene expression and modifies the identity of pharyngeal segments [116]. BRPF1 inactivation in mice leads to embryonic lethality at E9.5 and deactivation of cerebrum-specific BRPF1 in mice results in hypoplasia in the dentate gyrus. BRPF2 has also proven to play a critical role in erythropoiesis, and its absence in mice leads to embryonic lethality at E15.5, however, BRPF3 inactivation did not show any specific consequence in mice development [72, 117].

1.2.5. Potential inhibitors for KAT6A and BRPF1-related disease

The balance between histone acetylation, facilitated HATs, and histone deacetylation, mediated by HDACs, is typically tightly regulated. However, this balance is frequently disrupted in diseases like cancer [118]. In contrast to mutations in DNA sequence which are irreversible, chromatin modifications are mostly reversible. This presents opportunities for the utilization of

epigenetic therapy in clinical settings. The very first epigenetic drugs, Decitabine and Vorinostat, received FDA approval in 2006 and have since become accessible as cancer treatments medications for humans [119]. In previous research done in this lab hyper acetylation was reported in ten cancer cell lines, which contribute to the hypothesis of hyper acetylation may lead to cancer initiation and progression [120]. Therefore, utilizing inhibitors might be an effective therapy. The inhibitors are briefly reviewed below.

1.2.5.1 HDAC inhibitors

HDAC inhibitors have a high potential of functioning as a therapeutic drug [118]. Numerous HDAC inhibitors have been isolated from natural sources or artificially synthesized [118]. HDAC inhibitors can be classified into at least four structural categories: hydroxamates, cyclic peptides, aliphatic acids, and benzamides [118]. TSA (Trichostatin A) was the initial naturally occurring hydroxamate compound discovered to exhibit HDAC inhibitory activity [121]. Vorinostat, which shares a similar structure to TSA, was the pioneering HDAC inhibitor to receive FDA approval for the treatment of relapsed and refractory cutaneous T-cell lymphoma (CTCL) [122]. Among HDAC inhibitors, cyclic peptides represent the most intricate group in terms of structural complexity. This group encompasses compounds such as depsipeptide, apicidin, and a variety of cyclic peptides containing hydroxamic acid [123]. Aliphatic acids, including butyrate, and valproic acid, exhibit relatively modest HDAC inhibitory effects and require millimolar concentrations to exert their activity [124, 125]. SNDX-275, previously known as MS-275 is a synthetic compound derived from benzamide. It demonstrates inhibitory activity against HDAC1, HDAC2, and HDAC3 (belonging to class I HDACs) in the micromolar (μM) range [118]. Recent clinical trials have indicated that HDAC inhibitors hold promise as therapeutic agents, despite the fact that their specific targets and mechanisms of action are not fully understood. Moreover, the

potential for expanding their therapeutic applications beyond cancer treatment and also as a combination therapy with other therapeutics still requires more investigation [118].

1.2.5.2. KAT6A-specific inhibitors

KAT6A plays crucial roles in normal hematopoietic stem cells and is frequently targeted by chromosomal translocations, leading to the development of acute myeloid leukemia. Similarly, chromosomal translocations involving KAT6B have been detected in various types of cancers [78, 126]. KAT6A is involved in the suppression of cellular senescence by regulating the CDKN2A locus's suppressors, a function that relies on its KAT activity [127]. In mice with MYC-induced lymphoma, the survival rate significantly increases from 105 to 413 days when one allele of KAT6A is lost [128]. These findings suggest that inhibiting KAT6A and KAT6B could potentially provide therapeutic benefits for cancer treatment. As an example, WM-8014 and WM-1119 are biochemical compounds which inhibit the function of KAT6A and KAT6B respectively [129]. These compounds inhibit the KAT6A/B function by acting as reversible competitors of acetyl coenzyme A. While WM-8014 enhances oncogene-induced senescence in vitro, WM-1119 would prevent lymphoma progression in vivo [129].

1.2.5.3. BRPF1-specific inhibitors

The activity of bromodomain-containing proteins is commonly hindered by targeting their bromodomains. Multiple inhibitors specifically designed for bromodomains are accessible, and these inhibitors frequently interact with the Bromodomain and Extra-Terminal motif (BET) [130]. Significant progress has been made in the development of potent and specific inhibitors for the bromodomains of BET proteins (BRD2, BRD3, BRD4, BRDT). These inhibitors have shown promising results and several of them are currently being tested in clinical trials. While the majority of research has concentrated on the development of BET inhibitors, recent studies have revealed

the possibility of selectively targeting other bromodomains beyond the BET family [130].As a notable example, in a recent study GSK5959 which acts as a benzimidazolone probe to inhibit BRPF1, was used to reduce BRPF1 expression as a therapeutic approach for hepatocellular carcinoma (HCC). The use of GSK5959 to inhibit BRPF1 resulted in the suppression of colony formation and cell proliferation in various HCC cell lines that exhibited elevated levels of BRPF1 expression [131]. In addition, to examine the inhibitory effect in vivo, researchers administered GSK5959 to nude mice with tumors for a duration of two weeks [131]. The results indicated that GSK5959 effectively suppressed the growth of subcutaneous tumors in the mice, without any observable signs of adverse effects [131]. These findings suggest that GSK5959 holds promise as a potential therapeutic drug for the treatment of hepatocellular carcinoma (HCC) [131].

1.3 Rationale of the thesis project

The identification of over 500 epigenetic regulators via sequence analysis of the human genome has ushered in an era of exploration to address the questions about whether and how chromatin dysregulation influence a range of diseases [61]. While exome sequencing often uncovers various mutations in genes encoding these epigenetic regulators, our grasp of the pathological implications of these newly found mutations remains largely constrained.

My research project is centered around the goal to delve deeper into the mutations present in the genes encoding the paralogous epigenetic regulators BRPF1 and BRPF2, both of which play important roles in activating their respective lysine acetyltransferase complexes. As a multifaceted epigenetic regulator, BRPF1 is equipped with several chromatin-reading modules, including three dedicated to histone recognition and one for non-specific DNA-binding. Human BRPF2 and BRPF3 are paralogous to BRPF1, making these three form a unique subgroup within the extensive bromodomain superfamily that contains additional 39 members [61, 67, 117]. Prior research in our

and other laboratories unveiled the indispensable role of BRPF1 in mouse embryogenesis and forebrain formation [73, 132-134]. Concurrently, the loss of BRPF2 in mice culminates in embryonic lethality around E15.5 and thereby underscores its pivotal function in erythropoiesis [72]. Furthermore, studies on dozens of patients with monoallelic *BRPF1* mutations show that some of themaffect the PZP domain, resulting in reduced H3K23 acetylation [61, 120, 135]. One unclear but important issue is whether the PWWP domain of BRPF1 also plays a role in patients. Related to this, we have now identified 17 new patients with *BRPF1* mutations, with some of them affecting the PWWP domain, thereby highlighting the clinical importance of this domain (see Chapter 2). The importance of *BRPF1* mutations in patients with a new neurodevelopmental disorder raises the question about whether the same holds true for *BRPF2*. Related to this important question, we have also identified and investigated monoallelic *BRPF2* mutations in 8 patients (see Chapter 3). In conclusion, this project has identified and analyzed new *BRPF1* and *BRPF2* mutations, thereby strengthening their links to neurodevelopmental disorders.

Chapter 2 Role of BRPF1 in a neurodevelopmental disorder

2.1 Abstract

In this chapter, the detailed methodologies underpinning our investigation into BRPF1 mutations are thoroughly articulated, delineating our integrative approach combining genetic data acquisition from patients with laboratory replication techniques. The initial phase encompassed a rigorous collection of genetic data from patients, in collaboration with many clinicians (unpublished), aimed at identifying novel mutations within the BRPF1 gene. This endeavor resulted in the discovery of 17 previously undocumented mutations, several of which perturb the PWWP domain, underscoring the critical clinical relevance of this domain's integrity.With these genetic anomalies in hand, our next venture was to meticulously refine a site-directed mutagenesis protocol tailored to precisely emulate these patient-derived BRPF1 mutations in a controlled laboratory environment. We successfully generated all 17 BRPF1 variants, ensuring that each in-vitro mutant was an exact replica of the patient mutation, thereby setting the stage for robust downstream functional analyses.

Our fellow students are conducting histone acetyltransferase (HAT) activity assays, a vital component of our comprehensive investigation, to unravel the functional consequences of the BRPF1 mutations. Hopefully, these assays will provide invaluable insights into how these mutations may perturb the normal enzymatic functions that regulate chromatin structure and gene expression.

2.2 Introduction

In this chapter, we delve into the intricacies of our methodological advancements in the context of BRPF1 research, a gene that is pivotal for the formation and function of several histone

acetyltransferase (HAT) complexes, such as the MOZ/MORF complex. These complexes are instrumental in catalyzing the acetylation of histone lysine residues, a critical process for modulating chromatin structure and gene expression. The integrity of these complexes is vital for cellular homeostasis, and their dysfunction may precipitate a spectrum of disease states.

Our journey began with the challenge of optimizing the standard site-directed mutagenesis technique. After meticulous experimentation and refinement, we developed a superior version of this method, achieving an impressive efficiency of over 70%. This breakthrough enabled us to produce the desired BRPF1 mutations within a mere 48 hours from the initiation of the project. Following the generation of these mutations, we verified each one through Sanger sequencing, ensuring fidelity to the original patient-derived mutation.

With the mutations confirmed, we embarked on the arduous task of transfecting cells to express the mutant proteins. These proteins, once synthesized, underwent a purification process, allowing us to isolate them for further analysis. A critical step was the comparative expression analysis against the wild-type BRPF1 complex. This comparison was not merely a quality control checkpoint; it was a window into understanding how the mutations might perturb complex assembly. Imperfect expression profiles were our first indication of a potential mechanistic link between the mutations and the disease phenotype.

To probe the functional implications of these mutant complexes, we conducted histone acetyltransferase assays. These assays were pivotal; they compared the enzymatic activity of mutant complexes with that of the wild-type BRPF1. A decrease in HAT activity in the mutant complexes would signify a disruption of normal function, thus corroborating a connection between the mutation and disease pathology.

The advanced mutagenesis techniques and functional assays we have employed offer a promising avenue for dissecting the structural and functional ramifications of BRPF1 mutations. By understanding how these mutations influence complex assembly, histone acetylation patterns, and gene expression, we can elucidate the pathogenesis of related diseases. Furthermore, this knowledge affords us the opportunity to explore BRPF1 as a therapeutic target.

Our methodologies, which intersect state-of-the-art genetic editing with proteomic and genomic analyses, are designed to provide a comprehensive characterization of BRPF1's role in epigenetic regulation. This multi-pronged strategy aims to decode the precise molecular disruptions caused by BRPF1 mutations. Our collective efforts enhance our understanding of epigenetic modulators and pave the way for novel interventions in genetic disorders where BRPF1 is implicated.

Previous research, including studies from our laboratory, has underscored the significance of BRPF1 in developmental processes such as mouse embryogenesis and forebrain formation [73, 132-134]. In the clinical realm, the PZP domain of BRPF1 has been highlighted due to its mutations in patients [61, 120, 135]. Yet, the role of the PWWP domain remained elusive until our identification of 17 new patients with BRPF1 mutations impacting this domain, underlining its clinical relevance (Table 2.1).

2.3 Materials and Methods

2.3.1 Identification of BRPF1 mutations in patients

BRPF1 mutations in patients were identified by our clinical collaborators. The details will be described elsewhere (Table 1 and Fig. 2.5.2)).

2.3.2 Generation of Constructs

Previously, **the coding sequences for** BRPF1, ING5, and EAF6 were engineered onto pcDNA 3.1-HA [117]. KAT6A/B, were engineered onto pcDNA3.1-Flag [117]. Template DNA underwent PCR amplification and PFU high fidelity DNA polymerase (Agilent 600380). Methylated DNA fragments of PCR product undergo digestion using DpnI enzyme (NEB R0176). Subsequent transformation into DH5α and growth in liquid culture occurred, followed by plasmid purification (Qiagen 27106). Plasmids were sequenced using Sanger sequencing method for verification.

For the generation of mutants' variants PCR amplification with mutant specific primers were used. Primers were designed using a novel technique with significantly increased efficiency (Fig. 2.5.1) [136]. Based on the novel method, unique primers were designed for each mutant containing extended non-overlapping sequences at the 3' end and primer-primer complementary sequences at the 5' end [136].

Primers utilized to generate BRPF1 mutant variants are: R106H-F (5'-TTGCATGGCCACGTCCACCGCATCAGCATCTTTGACAACCTGGAT-3), R106H-R (5'-GCGGTGGACGtGGCCATGCAAGTCCACCTCCACCATGCGCTGGGC-3'), E208K-F (5'-TCTGCAGAGAAGCTGGACGAGGAAGTAGAGTATGACATGGACGAG-3') , E208K-R (5'- CGTCCAGCTTCTCTGCAGACTTCTCGATGTACCGGTAATAGGAAG-3'), R319C-F (5'-CCCCTATATCCCTGAGGGCCAGTGGCTGTGCCGCTGTTGCCTGCA-3'), R319C-R (5'-CAATCCACAGCACGAGAGGGTGACTGCAGGCAACAGCGGCACAGC-3'), V352L-F (5'- GGCCCATGTGCTGTGTGCCTTGTGGATCCCTGAGGTCTGC-3'), V352L-R (5'-AAGGCACACAGCACATGGGCCCAGCGCCCGTCATCTGTCT-3'), N364K-F (5'-CTTCGCCAAAACGGTCTTCCTAGAGCCTATTG-3'), N364K-R (5'-

GAAGACCGTTTTGGCGAAGCAGACCTCAGGGA-3'), I377V-F (5'-ATTGAGCACGTCCCACCAGCTCGCTGGAAGCTCACCTGCTACATT-3'), I377V-R (5'-CTGGTGGGACGTGCTCAATGCTGTCAATAGGCTCTAGGAAGACCG-3'), C386W-F (5'-AGCTCACCTGGTACATTTGCAAACAACGGGGCTCAGGGGCCTGCA-3'), C386W-R (5'-TGCAAATGTACCAGGTGAGCTTCCAGCGAGCTGGTGGGATGTGCT-3'), N445K-F (5'-GCCTACTGCAACATCCACACGCCTCCAGGTTC-3'), N445K-R (5'-TGTGGATGTTGCAGTAGGCTGTCTTGCGGACA-3'), A497T-F (5'-AATGAAGAAGACACGGAAGATCCTGGCAGAGAAGCGGGCAGCAGC-3'), A497T-R (5'- ATCTTCCGTGTCTTCTTCATTTTGATCCGGGACTTGGCCTTGGCC-3'), N615S-F (5'-GACACAGGCAGCATCTTCAGCGAGCCGGTCCCTCTGTCTGAGGTA-3'), N615S-R (5'-GCTGAAGATGCTGCCTGTGTCCTTCTCTTGGAGCTGCTCCAAGGT-3'), A815E-F (5'-CACGGCGTGAAAAGATGATCAAGAAAGAGATG-3'), A815E-R (5'-ATCATCTTTtCACGCCGTGAGCGGCCCACACT -3'), V1095L-F (5'-TCTGGACCTCcTGTGGGCCAAATGCCGAGGCTATCCATCA-3'), V1095L-R (5'-TTGGCCCACAgGAGGTCCAGAGCATCCAGCGGGGAGTCCT-3'), P1106T-F (5'-CGTGTGGGCCAAATGCCGAGGCTATCCATCATACACAGCTCTGAT-3'), P1106S-R (5'-TCTCGGGGCATCTTTGGATCAATGATCAGAGCTGAGTATGATGGA-3'), P1130L-F (5'-TGTTCCACCATGGGGTTCCCATCCCTGTGCCCCTACTGGAGGTGC-3'), P1130L-R (5'-GTCATCTGCTCCCCAAGTTTCAGCACCTCCAGTAGGGGCACAGGG-3'), F1154del-F (5'-TCGTCCTCTTCGACAACAAACGAACCTGGCAGTGGCTGCCCAGGA-3'), F1154del-R (5'-CGTTTGTTGTCGAAGAGGACGAGGTAGAGATGCTCTCGGGCTTCC-3'), R1158*-F (5'- TGACAACAAAtGAACCTGGCAGTGGCTGCCCAGGACCAAGCTGGT-3'), R1158*-R (5'- TGCCAGGTTCATTTGTTGTCAAAGAAGAGGACGAGGTAGAGATGC-3'), R1191C-F

(5’-GTCCAACATCTGCAAGTCAGTACAGATCGCCTACCACAGGGCTCT-3’), R1191C-R (5’-ACTGACTTGCAGATGTTGGACTTGCGGCCCTCCAGCATCTTCTCC-3’), X1221G-F (5’- CGATAGTGATGGATACTGCTCAACACAGCCCAACCTATAG-3’), and X1221G-R (5’-GAGCAGTATCCATCACTATCGCTGGTCTCACTGCTCTGCT-3’).

2.3.3 Cell Culture and Transfections

HEK293 (human embryonic kidney 293) cells were cultured and passaged in Dulbecco’s modified Eagle’s medium (DMEM, Gibco) with 10% heat-inactive fetal bovine serum (FBS, Sigma), and 100 units/ml penicillin (P/S, Gibco). The cells were maintained in 37°C under 5% CO2.

In order to determine the effect of mutations on protein expression, HEK293 cells were transfected. On the day before transfection, cells were seeded at $2x10^6$ cells per 10 cm dish. 15 µl Lipofectamine 2000 (Thermofisher 11668019) was used to transfect 10 µg of plasmid DNA into one dish of cells. More specifically, to expressBRPF1 and its mutants, 3µg of a vector expressing HA-tagged BRPF1 (or its mutants), 5 µg of the vector for FLAG-tagged KAT6,1 µg of the vector for HA-tagged ING5 and 1 µg of the vector for HA-tagged EAF6 were transfected into one dish of cells. Plasmids and 18 µl of Lipofectamine 2000 were mixed in antibiotic-free DMEM/FBS media. The mixture was added to one dish of cells for further incubation at 37°C under 5% CO2 for about 5 hours. Afterwards, the media was changed to DMEM with FBS and P/S (Penicillin-Streptomycin). 48 hours post-transfection, the media were aspirated and cells were washed with PBS for for further analysis, as detailed below.

2.3.4 Immunoprecipitation and Acetylation Assays

Forty-eight hours post transfection, cells were washed twice with PBS and soluble protein extracts were obtained for affinity purification on M2 agarose, as previously described by Yan *et al*. [137]. FLAG peptide (Sigma F3290) was used to elute bound proteins from M2 agarose beads (Millipore Sigma A2220). Both whole cell extracts and affinity purified were prepared for immunoblotting to detect complex formation with anti-FLAG (Sigma F3165) and anti-HA (Biolegend), as previously described by Yan *et al.* [137]. Post primary incubation, membranes were washed with TBST and afterward incubated with secondary antibodies. After washing membrane once more and incubating the membrane with Substrate Super Signal West Pico Plus Chemiluminescent Substrate, the immunoblotting signals were developed.

Histone acetylation assays were performed based on modified protocol by Yan *et al.* [137]. The purified protein complexes of acetyltransferase were subjected to acetylation reactions using a mixture of nano-pure water, histone or nucleosome substrates, acetyl-CoA, and 5x buffer AQ. After incubating the reactions at 37 °C for 1 hour, the reactions were stopped and the samples were used for SDS-PAGE and subsequent immunoblotting withanti-H3K14ac (EMD Millipore 07-353), anti-H3K23ac (EMD Millipore 07-355), anti-H4 (Abcam ab18253), anti-H4K5ac (Millipore Sigma 07-327), anti-H4K8ac (Millipore Sigma 07-328) antibodies, as previously described by Yan *et al.* [137].

2.4 Results

2.4.1 Optimizing a site-drected mutagenesis method

Site-directed mutagenesis serves as a fundamental tool in contemporary molecular biology, facilitating precise manipulation of protein sequences. This method is indispensable in the exploration of function, as well as in the fields of genetic modification, biochemistry, and protein

engineering [138]. One site-directed mutagenesis is known as the Quick-exchnage mutagenesis method (Fig. 2.4.1A), which is based on PCR with a pair of primers and subsequent DpnI digestion to remove the parent plasmid. While the technique in question has demonstrated considerable efficacy in experiments carried out by previous trainees in our laboratory, it took considerable efforts to generate a mutant (often requiring multiple trials. Related to thismy own initial attempts to apply this Quick-exchnage mutagenesis method to genertating the new BRPF1 mutants (Table 1). BRPF1 were met with significant challenges in efficiently producing mutants.

Despite being highly beneficial and straightforward (as well as its wide use in different laboratories), the QuickChange™ mutagenesis method (Agilent) presents certain constraints. Due to the primers' complete overlapping, there is a propensity for self-annealing, which necessitates meticulous primer design to prevent self-pairing that can interfere with the annealing to the template. Moreover, as the newly synthesized DNA strands are 'nicked,' they are unsuitable for use as templates for further amplification, unlike in conventional PCR.

To overcome these drawbacks, we have sought an alternative method. Related to this, we noticed Liu *et al.* [136] has developed a new mthod employing primers with elongated non-overlapping regions at the 3' end, which are considerably longer than those recommended in the reference, and sequences at the 5' end that are complementary to each other (Fig. 2.4.1B). We have thus investigated whethere this relatively new method works better than the Quick-exchange method.

As a result, we have tested a new method by adopting the pimer design strategy employed by by Liu *et al.* (Fig. 2.5.1B) [136]. This protocol retains the straightforward single-step procedure characteristic of the QuikChange™ site-directed mutagenesis approach, but it enhances its efficiency via a different primer design strategy. This method incorporates an innovative primer

design strategy, which not only efficiently prevents primer dimerization (thereby improving primer-template annealing) but also allows the newly synthesized DNA strands to serve as templates for successive rounds of amplification (Fig. 2.5.1B) [136].

To improve this method further, we substantially reduced amount of parental DNA (from 20-50 ng to 10 ng) and minimzing the reaction volume to 10 ul. With this optimized method, we achieved a high effieicny of 50-100%. The lower amount of template DNA simplifies DpnI digestion following PCR amplification, thereby enhancing the general efficacy and reproducibility of this method. The outcomes revealed that, while not incurring extra reagent expenses beyond those associated with the QuikChange™ method, our optimized protocol elevates the success rate significantly, reaching the efficieny 50-100% (Table 2.2). Typically, per mutation, we only need to send plasmids from three colonies for Sanger sequencing so we can generate 5-10 different mutants easily within a week. By comparison, former trainees in this laboratory and also myself needed to spend weeks (sometimes months) of trials and errors, for optimizing PCR conditions, to obtain a few mutants when we used the Quick-exchange mutagenesis mehod (Mousavi, N. & Yang, X.J., manuscript under preparation).

Furthermore, we also tested different primers and found primers shorter than those used by Liu *et al*. are also OK (Fig. 2.5.1C), thereby reducing the cost and minimize the possibility of unwanted mutations introduced by premature primer molecules (this is because during primer synthesis, each cycle of nucleotide addition has an efficiency slightly lower than 100%). Due to chemical synthesis, shorter primers typically have higher quality and thus reduce unwanated mutations introduced by some incorrect molecules as impurity in unpurified primers. Overall, our optimized method has improved the currently repported mutagenesis methods that are being widely used in different laboratories. Using this optimized method, we have generated 17 new

BRPF1 mutants, whose sequences were all confimed by Sanger sequencing (Fig. 2.5.5). A short manuscript on this improved method is being prepared for submission (Mousavi, N. & Yang, X.J., unpublished).

2.4.2 Generation and analysis of 17 BRPF1 mutants derived from patients

Through international collaboration with multiple physicians (unpublished data), we have identified 17 patients with Intellectual Disability via Exome Sequencing. Every patient was directed to our laboratory because their WES findings revealed they carry mutations in the BRPF1 gene. These subjects are all from different families and they are *de novo* mutations. To determine if these variants have the potential to cause the Intellectual Disability phenotype they must be analyzed. All the mutants can be found in Table 2.1.

Among the *BRPF1* mutations, fifteen are missense and the remaining two are nonsense mutations or reading frameshifts that lead to C-terminal truncations of the protein (Table 2.1). Mutation sites are on various regions of BRPF1. The mutants which the ING5- and MEAF6-interacting domain is intact in them are expected to form tetrameric complexes with KAT6A (or KAT6B), ING5, and MEAF6. Hence, the variations seem to create unique sets of variants, indicating that these mutations could disrupt BRPF1 functions via various methods.

In understanding that *BRPF1* mutations likely impact function, structure, or stability of *BRPF1*, it is possible that they are responsible for causing the clinical features found in the patients, thus we decided to proceed with further analysis of the mutations.

KAT6 FLAG-tagged plasmids were generated through PCR amplification of two KAT6 fragments with restriction sites, standard cloning procedures were then followed. The insertion of the full KAT6 gene into the FLAG vector was confirmed by colony PCR, restriction digestion and

validated by sequencing. This process was repeated to generate HA-tagged plasmids of the MYST complex, namely, BRPF1, ING5 and EAF6. Mutations in the BRPF1 cDNA were generated via site directed mutagenesis, with primers designed to specifically insert the patient mutation. Mutation insertion was confirmed with sequencing.

The MYST complex plasmids, along with the BRPF1 mutants were used in the following experiments to test the differences in BRPF1. To begin assessing individual differences between the wild type and mutant BRPF1 proteins a transfection of *BRPF1* was performed. For this analysis BRPF1 plasmids were transfected into HEK293 cells and purified through co-IP. The PWWP domain variants (Fig. 2.5.6) were predicted to be unable to form tetrameric complexes with KAT6A (or KAT6B), ING5, and MEAF6, however the results indicate that the mutants were able to form tetrameric complexes with KAT6A, ING5, and MEAF6. The rest of the BRPF1 mutants also were able to form the tetrameric complex, expect for p.Arg106His (c.318G>A) and p.Glu208Lys (c.622G>A) that were unable to form the tetrameric complex. The histone acetyl transferase assays are being carried by others in the lab to further investigate the functionality of the variants (Fig. 2.5.6).

In conclusion, we have identified 17 new patients with *BRPF1* mutations, with some of them affecting the PWWP domain, thereby highlighting the clinical importance of this domain. We have optimized a mutagenesis method and obtained all these 17 mutants, thereby setting the stage for futher functional validation as being carried now in our laboratory (Table 2.1 and Fig. 2.5.2).

2.5 Illustrations

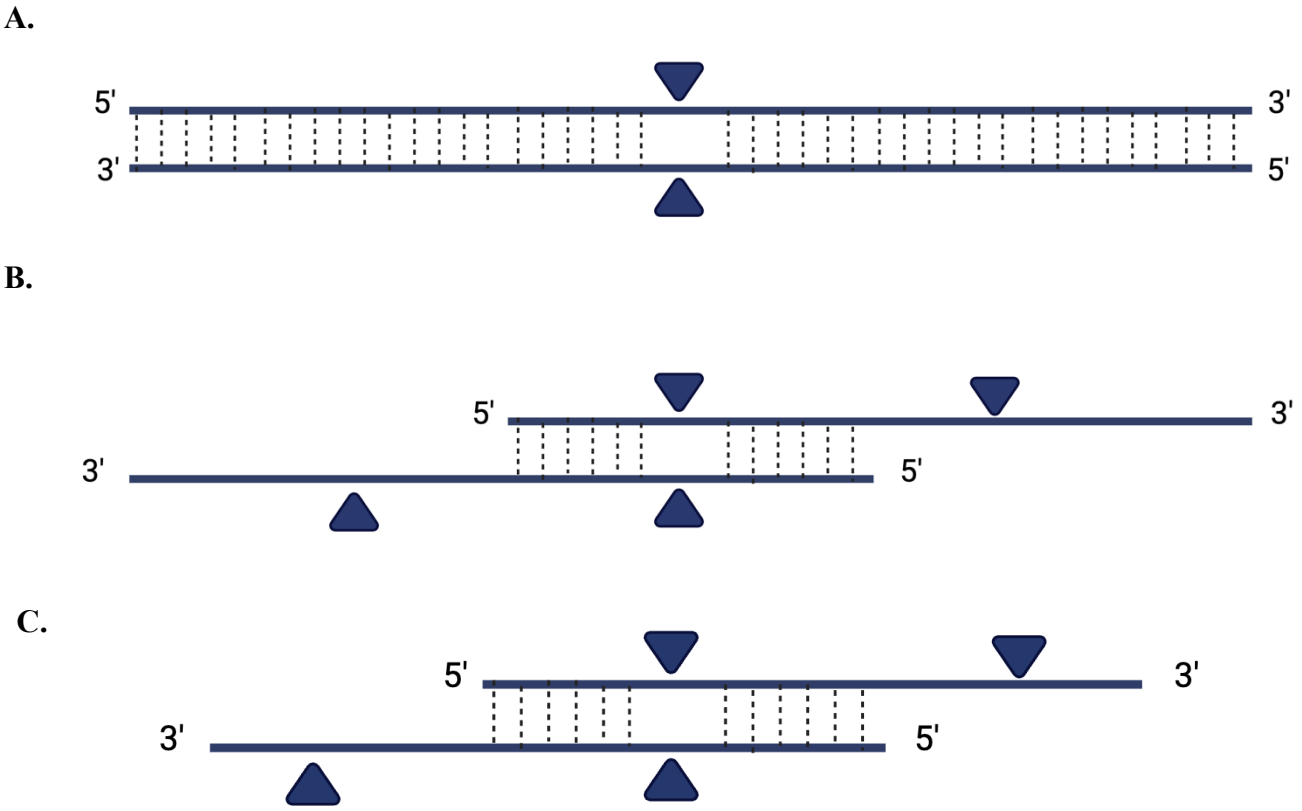
Table 2.1 Identification of *BRPF1* Mutations in Seventeen Individuals

Individual	Mutation (GenBank: NM_001003694.2)	Substitution in BRPF1
P1	c.318G>A	p.Arg106His
P2	c.622G>A	p.Glu208Lys
P3	c.955C>T	p.Arg319Cys
P4	c.1056G>C	p.Val352Leu
P5	c.1089G>A	p.Asp364Lys
P6	c.1129A>G	p.Ile377Val
P7	c.1158C>G	p.Cys386Trp
P8	c.1335G>C	p.Asn445Lys
P9	c.1489G>A	p.Ala497Thr
P10	c.1844A>G	p.Asn615Ser
P11	c.3283G>C	p.Val1095Leu
P12	c.3316C>A / c.3316C>T	p.Pro1106Thr/Ser
P13	c.3389C>T	p.Pro1130Leu
P14	c.3461T>G	p.Phe1154del*
P15	c.3472C>T	p.Arg1158*
P16	c.3571C>T	p.Arg1191Cys
P17	c.3661T>G	p.X1221Gly

Table 2.2 Efficiency of the optimized site-directed mutagenesis method for generating representative BRPF1 Mutations

Mutation	Number of positive mutants sequenced
p.Arg106His (c.318G>A)	**5 out of 5 (100%)**
p.Glu208Lys (c.622G>A)	**4 out of 5 (80%)**
p.Asp364Lys (c.1089G>A)	**3 out of 4 (75%)**
p.Cys386Trp (c.1158C>G)	**3 out of 3 (100%)**
p.Ala497Thr (c.1489G>A)	**5 out of 6 (83.3%)**
p.Asn615Ser (c.1844A>G)	**3 out of 4 (75%)**
p.Phe1154del* (c.3461T>G)	**4 out of 4 (100%)**
p.X1221Gly (c.3661T>G)	**2 out of 3 (67%)**

Figure 2.5.1



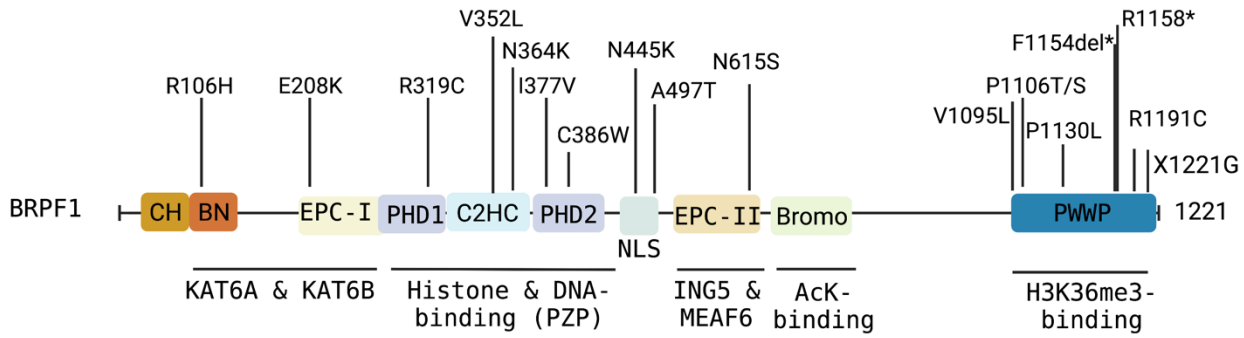
Schematic showing the primer design for three different site-directed mutagenesis methods

A. The original primer designs are shown for site-directed mutation, with triangles marking the mutation sites within the primer sequences. The primers are **20-25 nucleotides in length**. This is based on the Quick-echange mutagenesis method.

B. The optimized primer designs are shown for site-directed mutation, with triangles marking the mutation sites within the primer sequences.Liu *et al.* designed primers with about **45 nucleotide**s in length, which was our initial design as well. The illustration is adapted from Liu *et al.* 2008 paper [138].

C. The optimized primer designs are shown for site-directed mutation, with triangles marking the mutation sites within the primer sequences. We have optimized the primer design by reducing the length of the primers to **30 nucleotides in length**, which is much shorter than that in panel B.

Figure 2.5.2

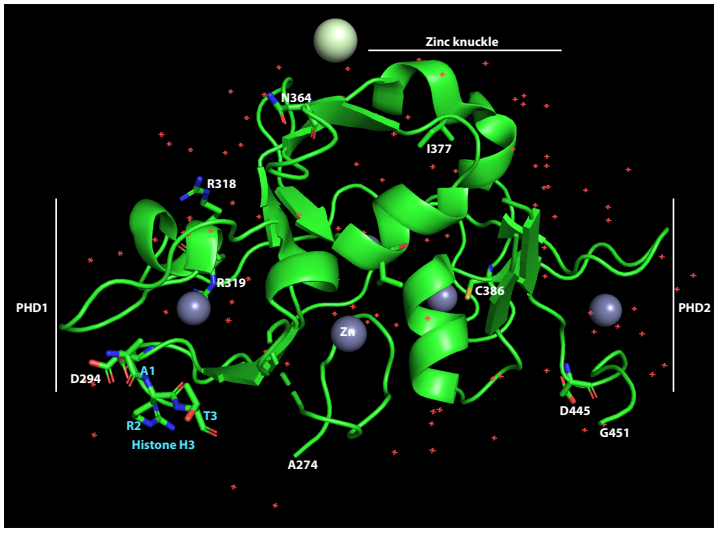


Location of new BRPF1 mutations

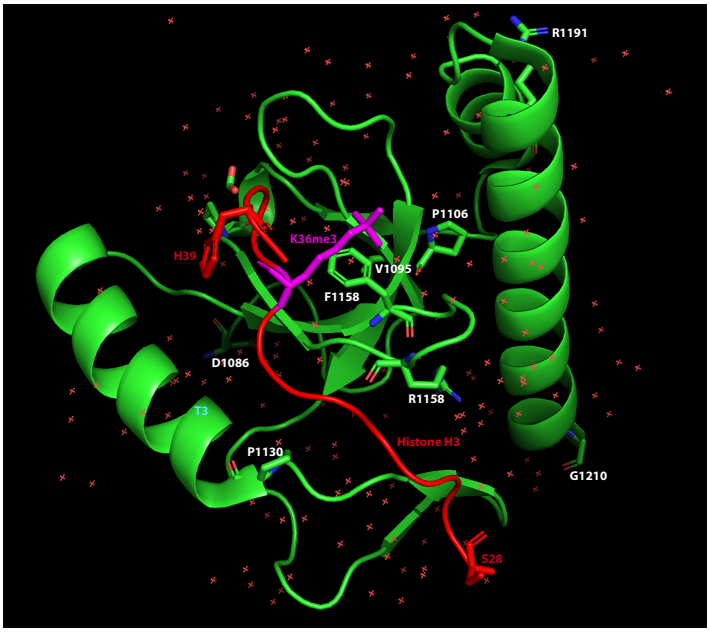
The schematic representation of BRPF1 along with seventeen allelic variants identified in a corresponding number of individuals. The precise DNA sequence alterations are catalogued in Table 2.1. The BRPF1 protein is composed of several functional modules critical for chromatin engagement, including the PZP domain, a bromodomain, and a PWWP domain. The PZP domain is an assemblage of two plant homeodomain (PHD) fingers contiguous with a C2HC zinc finger, with the former PHD finger having an affinity for the N-terminal end of histone H3. The amalgamation of the C2HC zinc finger and the subsequent PHD finger constitutes a domain with a general DNA-binding capability. The bromodomain possesses the capacity to bind acetyllysine residues, whereas the PWWP domain specifically interacts with trimethylated histone H3. A pivotal EPC-like motif situated at the C-terminus adjacent to the PZP domain is indispensable for the formation of a stable trimeric complex with the proteins ING5 and MEAF6. Furthermore, through the EPC-like motif located at the N-terminal side proximal to the PZP domain, along with an additional conserved region more proximal to the N-terminus relative to this motif, BRPF1 engages and potentiates the enzymatic activity of KAT6A, KAT6B, and KAT7. Image generated with BioRender.com.

Figure 2.5.3

A.



B.

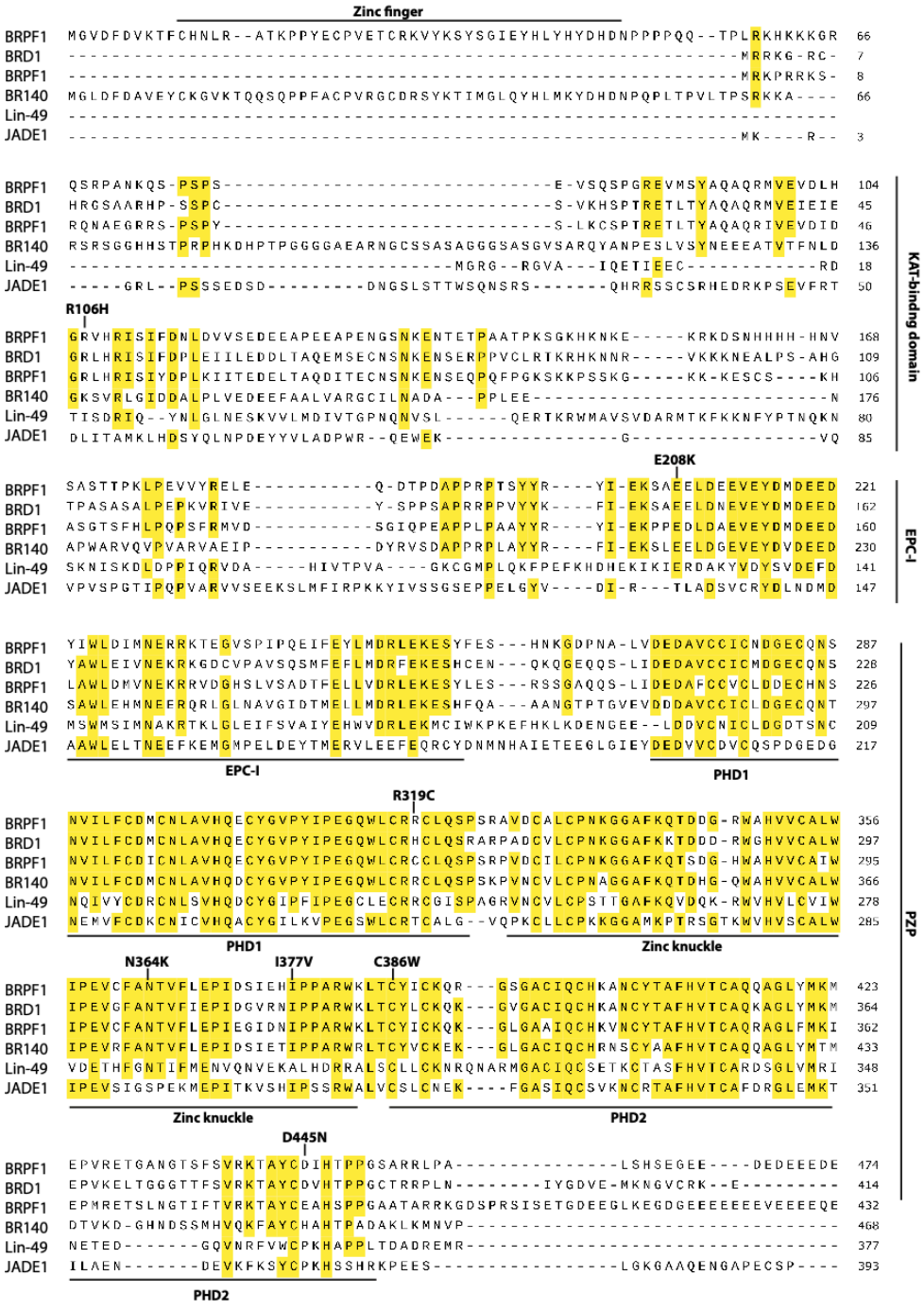


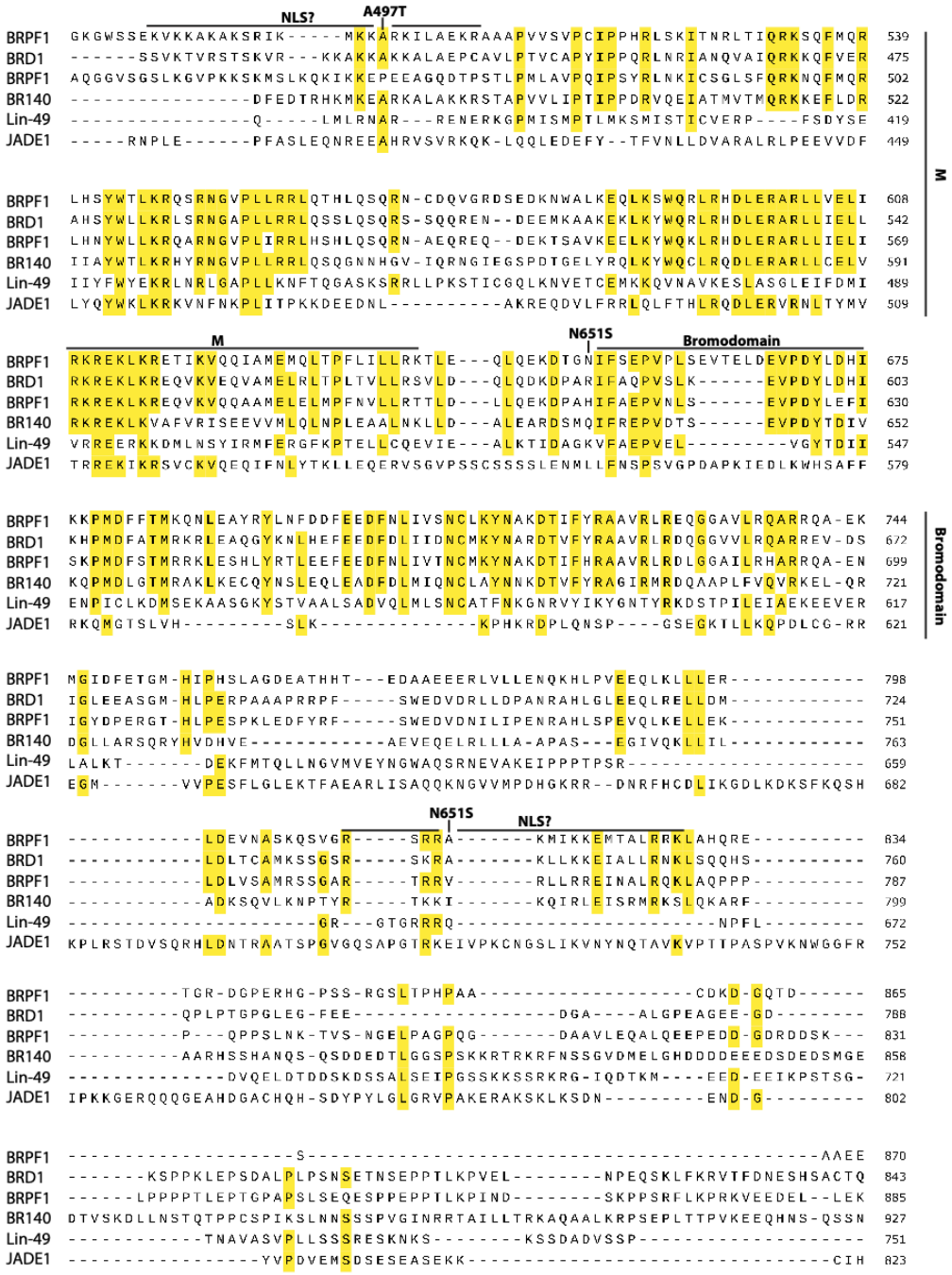
Structural model showing the effect of BRPF1 mutations on its structure

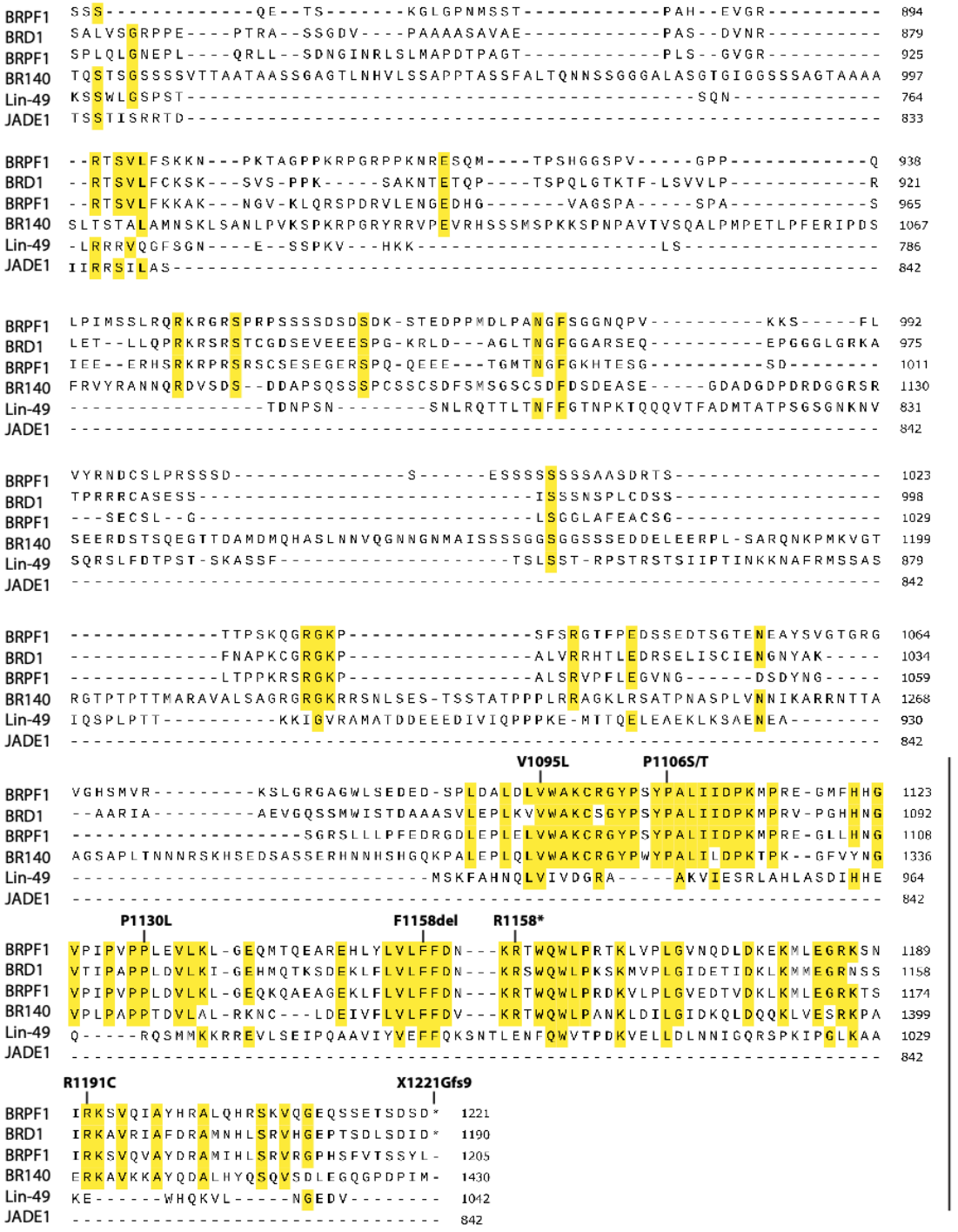
(A) Non-PWWP domain mutations: Schematic representation of the effect of various BRPF1 mutations on non-PWWP domain. Highlighted are specific amino acid residues where mutations have occurred, demonstrating alterations in the protein's three-dimensional conformation. Each mutation is annotated to show its position relative to functional sites such as the zinc knuckle and interaction interfaces with other proteins or DNA. This model, generated using PyMOL, provides insight into how these non-PWWP domain mutations might disrupt BRPF1's normal function, potentially leading to aberrant chromatin states and gene expression profiles.

(B) PWWP domain: This panel focuses on the PWWP domain, known for its role in binding to methylated histone tails and mediating protein-protein interactions crucial for chromatin organization. The schematic highlights the BRPF1 PWWP domain's structure, emphasizing the specific mutations and their proximity to critical regions such as the histone H3 binding site. The structural model, also generated with PyMOL, underscores the potential for these mutations to influence gene regulatory mechanisms and the importance of the PWWP domain in maintaining proper chromatin architecture.

Figure 2.5.4


Zinc finger
BRPF1
BRD1
BRPF1
BR140
Lin-49
JADE1
KAT-binding domain
R106H
E208K
EPC-I
PHD1
Zinc knuckle
R319C
N364K
I377V
C386W
PZP
PHD2
D445N

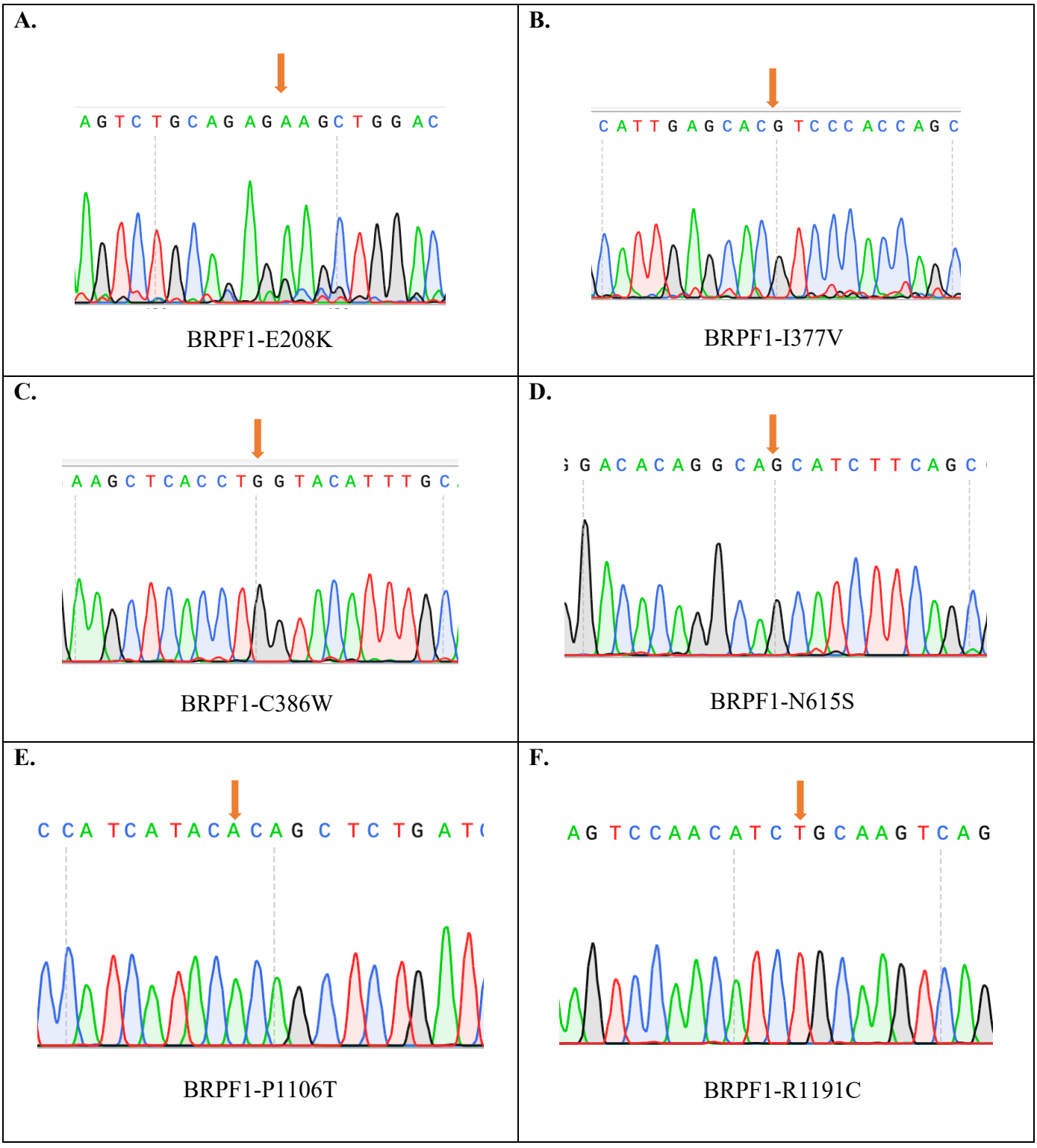

NLS?
A497T
M
N651S
Bromodomain
M
Bromodomain
N651S
NLS?
BRPF1
BRD1
BRPF1
BR140
Lin-49
JADE1


BRPF1
BRD1
BRPF1
BR140
Lin-49
JADE1
V1095L
P1106S/T
P1130L
F1158del
R1158*
R1191C
X1221Gfs9
PWWP

Multiple sequence alignment highlighting conserved regions and mutation sites in BRPF1

The sequence of human BRPF1 is aligned with its paralog BRD1, and orthologs from C. elegans (Lin-49), JADE1, and BR140, showcasing evolutionary conservation across species. Highlighted regions indicate zinc finger domains, bromodomains, PWWP domains, and PHD fingers, which are characteristic features of the protein involved in chromatin interaction and recognition. Specific mutations observed in individuals P2, P4, P5, P6, P11, and P14 are marked and correspond to codons that are conserved across a range from Drosophila to humans, indicating their potential critical role in the protein's function. The mutations lead to changes in amino acids that are essential for the structural and functional integrity of BRPF1, suggesting a link to the associated phenotype. Yellow shading highlights conserved residues, while black text signifies the location of the mutation sites in the human sequence.

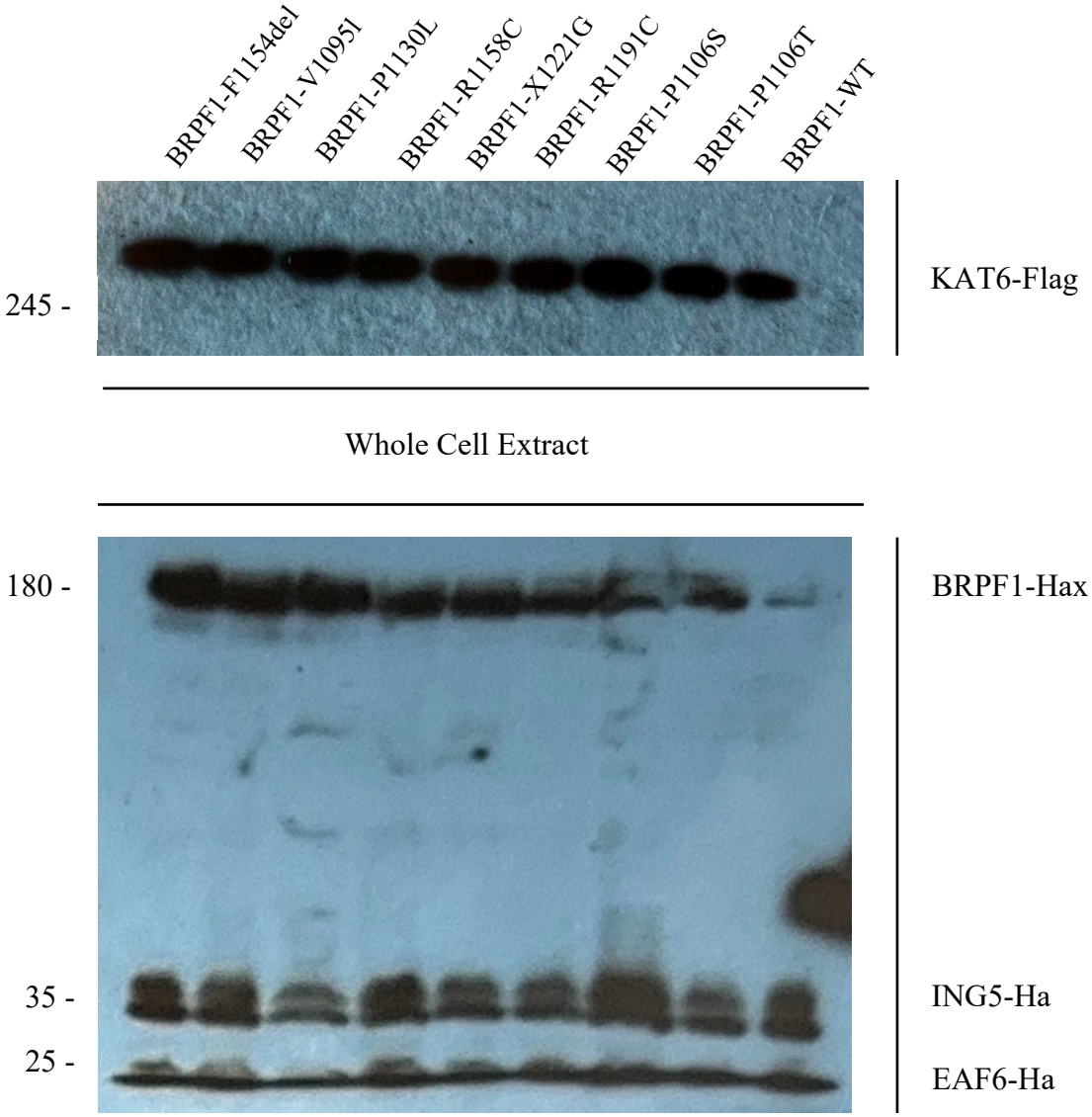
Figure 2.5.5

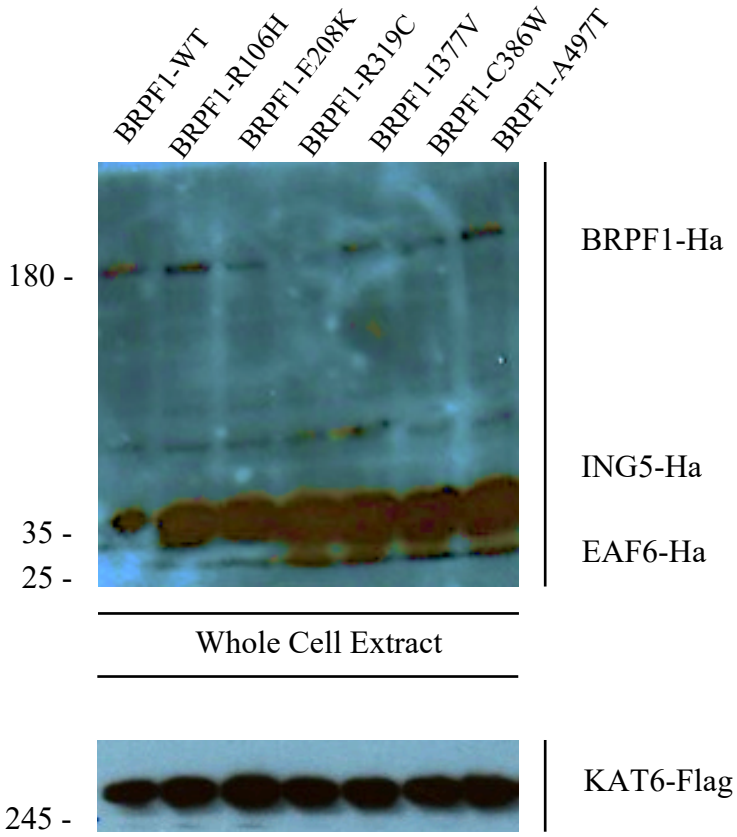

A.
A G T C T G C A G A G A A G C T G G A C
BRPF1-E208K
B.
C A T T G A G C A C G T C C C A C C A G C
BRPF1-I377V
C.
A A G C T C A C C T G G T A C A T T T G C
BRPF1-C386W
D.
G A C A C A G G C A G C A T C T T C A G C
BRPF1-N615S
E.
C C A T C A T A C A C A G C T C T G A T
BRPF1-P1106T
F.
A G T C C A A C A T C T G C A A G T C A G
BRPF1-R1191C

Representative Sanger Sequencing Results

Sanger sequencing chromatograms showing site-directed mutagenesis of BRPF1 gene at specific locations. Each panel represents a different mutation within the BRPF1 gene, indicated by the arrows, which confirm the introduction of the intended point mutations. Panel A exhibits the BRPF1-E208K mutation with a clear peak change at the mutation site. Panel B shows the BRPF1-I377V mutation, Panel C the BRPF1-C386W mutation, Panel D the BRPF1-N615S mutation, Panel E the BRPF1-P1106T mutation, and Panel F the BRPF1-R1191C mutation. Each chromatogram displays a single nucleotide substitution corresponding to the respective amino acid change in the BRPF1 protein. The sequence data illustrate the precision of the mutagenesis technique, evidenced by the singular peak alterations at the targeted positions without additional unintended mutations in the surrounding sequence. Chromatograms were analysed using SnapGene.

Figure 2.5.6


BRPF1-F1154del
BRPF1-V1095I
BRPF1-P1130L
BRPF1-R1158C
BRPF1-X1221G
BRPF1-R1191C
BRPF1-P1106S
BRPF1-P1106T
BRPF1-WT
245 -
KAT6-Flag
Whole Cell Extract
180 -
BRPF1-Hax
35 -
ING5-Ha
25 -
EAF6-Ha



BRPF1 complex formation with KAT6, ING5 and EAF6

Functional Characterization of BRPF1 Variants In Vitro (A) Interaction of BRPF1 and the variants with KAT6A, ING5, and MEAF6. KAT6A was produced in HEK293 cells as a FLAG-tagged fusion protein along with HA-tagged BRPF1 (or other variants), ING5, and MEAF6 as indicated. Soluble protein extracts were prepared for affinity purification on anti-FLAG agarose, and bound proteins were eluted with the FLAG peptide for immunoblotting with anti-FLAG and -HA antibodies.

Chapter 3 Role of BRPF2 in a new neurodevelopmental disorder

3.1 Abstract

BRPF1 mutations have been linked to reduced H3K23 acetylation, a specific modification on histones that can affect gene expression. This has been observed in patients with a new neurodevelopmental disorder. The identification of new patients with BRPF1 mutations affecting the PWWP domain (published) and the PWWP domain (see Chapter II) underscores the significance of these domains in clinical outcomes. Thisraises the question about whether *BRPF2* mutations could also lead to developmental anomalies, considering its important role in mouse development. The research in this chapter has identified *BRPF2* mutations in patients, suggesting a link to a new neurodevelopmental disorder, although this connection seems to be at an early stage of investigation. Thischapter also discusses gene functions, the impact of genetic mutations on development, and the association of these mutations with human diseases. It emphasizes the importance of understanding these gene functions and mutations for developmental biology and medicine.

3.2 Introduction

BRPF2 and BRPF1 are described as paralogous, meaning they are genes that have evolved by duplication within a genome and have evolved to carry out different functions. Despite their similarity in amino acid sequences, the divergence in their roles indicates that they have undergone a functional specialization after the duplication event. The molecular function of a gene product refers to the biochemical activity it participates in. BRPF2 and BRPF1 have different preferences in forming complexes with different KAT (lysine acetyltransferase) enzymes. BRPF2 prefers forming complexes with KAT7, while BRPF1 targets KAT6A and KAT6B. These interactions are

crucial as they suggest that BRPF1 and BRPF2 might regulate different sets of genes due to their association with different KAT enzymes, which are known to modify histones and thereby affect gene expression. The absence of BRPF2 in mice leads to embryonic lethality, which highlights its essential role in embryogenesis and erythropoiesis (the production of red blood cells). This lethality indicates that BRPF2 is critical for the proper development and survival of the embryo.

In this chapter, we delve deeply into the methodology used in our investigation of BRPF2, a critical component of epigenetic regulation via histone acetyltransferase (HAT) complexes. Our research primarily focused on the optimization of mutagenesis techniques, an approach that proved highly effective in our study, achieving over 70% efficacy. This success mirrors our previous work with BRPF1 mutants, where we encountered significant challenges using standard mutagenesis methods.

Our investigative journey began with the identification of the initial mutation sites in BRPF2. We then employed an optimized site-directed mutagenesis approach to create mutants in the laboratory. These mutants were subsequently verified through Sanger sequencing, ensuring the accuracy of our mutagenesis process. Following the confirmation of these mutations, we initiated transfection experiments to produce proteins derived from the mutated genes. The next critical step involved the purification of these proteins. We conducted a rigorous comparison between the expression levels of the mutated proteins and the Wild Type Complex, focusing on four key components: BRPF2, KAT7, ING5, and EAF6. This comparison was crucial to determine whether the proteins were well-expressed. Instances of imperfect expression were indicative of a potential correlation between the mutations and disease pathogenesis.

To further elucidate the functional implications of these mutations, we conducted a Histone Acetyltransferase (HAT) assay. This assay allowed us to compare the functionality of the mutant

complexes with that of the BRPF2 Wild Type. A decrease in functionality in the mutant complexes would suggest a direct link between the mutations and the disease. Conversely, if the functionality remained comparable to the Wild Type, it would indicate no direct correlation between the mutations and the disease.

Human BRPF2 and BRPF3 are paralogous to BRPF1, making these three form a unique subgroup within the extensive bromodomain superfamily that contains additional 39 members [61, 67, 117]. The absence of BRPF2 in mice culminates in embryonic lethality around E15.5 and underscores its pivotal function in erythropoiesis [72]. The importance of *BRPF1* mutations in patients with a new developmental disorder raises the intriguing possibility that *BRPF2* is also mutated in patients. Related to this, we have now identified and investigated monoallelic *BRPF2* mutations in 8 patients. This is entirely new as no reports on BRPF2 in a developmental disorder has been made. Thus, this project has made the first direct link of *BRPF2* mutations to a neurodevelopmental disorder.

3.3 Materials and Methods

3.3.1 Identification of BRPF2 mutations in patients

Subjects carrying BRPF2 were identified by our clinical collaborators (unpublished data). The details will be described elsewhere (Table 3.1 and Fig. 3.5.1).

3.3.2 Generation of Constructs

Previously, BRPF2, ING5, and EAF6 were engineered onto pcDNA 3.1-HA. HBO1 is engineered onto pcDNA3.1-flag [117]. Template DNA underwent PCR amplification and PFU high fidelity DNA polymerase (Agilent 600380). Methylated DNA fragments of PCR product undergo digestion using DpnI digestion enzyme (NEB R0176). Subsequent transformation into

DH5α and growth in liquid culture occurred, followed by plasmid purification (Qiagen 27106). Plasmids were sequenced using Sanger sequencing method for verification (Fig. 3.5.3).

For the generation of mutants' variants PCR amplification with mutant specific primers were used. Primers were designed using a novel technique with significantly increased efficiency [136]. Based on the novel method, unique primers were designed for each mutant containing extended non-overlapping sequences at the 3' end and primer-primer complementary sequences at the 5' end [136].

Primers utilized to generate mutant variants are: for BRPF2 C327R-F (5'-GAGGAACATCCCTCCAGCCCGGTGGAAACTGACACGCTACCTCTG-3'), C327R-R(5'-CAGGCACCCACGCCCTTCTGCTTACAGAGGTAGCGTGTCAGTTTC-3'), V379I-F (5'-CACCTTCTCCaTCAGAAAGACCGCTTACTGTGATGTCCACACGCC-3'), V379I-R (5'-GTCTTTCTGAtGGAGAAGGTGGTGCCACCGCCAGTCAGTTCCTTC-3'), R504*-F (5-CAGTCTCAGTGAAGCTCACAGCAGAGAGAAAATGATGAGGAGAT-3'), R504*-R (5'-GTGAGCTTCACTGAGACTGCAGGCTGGACTGCAGCCGCCGCAGC-3'), R534C-F (5'-CGACCTGGAGtGCGCTCGCCTGCTGATCGAGCTGCTGCGCAAGCG-3'), R534C-R (5'-AGGCGAGCGCaCTCCAGGTCGTGCCGCAGCCGCTGCCAGTACTTC-3'), L683Rfs-F (5'-GGGATGCACCGCCTGAGCGGCCTGCTGCGGCACCGCGGCGGCCT-3'), L683Rfs-R (5'-CGCTCAGGCGGTGCATCCCCGAGGCCTCTTCCAAGCCGATGCT-3'), Q761H-F (5'-TTCTCCGAAACAAGCTGAGCCAGCAGCACAGCCACCCCCTGCCCA-3'), Q761H-R (5'-TTCTCCGAAACAAGCTGAGCCAGCAGCACAGCCAcCCCCTGCCCA-3'), R931G-F2 (5'-GCAGCCAcGtAAAgGGTCGCGGAGCACATGCGGAGACTCCGAGGTGGA-3'), R931G-R (5'-CTCCGCGACCcTTTCCTTGGCTGCAGAAGAGTCTCCAACCTCGGA-3'), , P1133R -F (5'-AGTGGCTTCgTAAGTCCAAAATGGTTCCCCTT-3'), P1133R-R (5'-

TTGGACTTAcGAAGCCACTGCCAACTTCTCTT-3'), I1188F-F (5'-CTCAGTGACtTTGACTGACGGCCCGGCCGCCA-3'), and I1188F-R (5'-GTCAGTCAAaGTCACTGAGGTCGCTGGTCGGC-3'). It is noteworthy, that R931G-F1 did not function as expected due to an error in the primer sequence design, therefore, we designed R931G-F2 primer.

3.3.3 Cell Culture and Transfections

HEK293 (Human embryonic kidney 293) cells were cultured and passaged in Dulbecco's modified Eagle's medium (DMEM, Gibco) with 10% heat-inactive fetal bovine serum (FBS, Sigma), and 100 units/ml penicillin (P/S, Gibco) and were maintained in 37°C under 5% CO2.

In order to determine the effect of mutations on protein expression, HEK293 cells were transfected. The day before transfection, cells were seeded at $2x10^6$ cells per 10cm dish. Between 1.5-2µl Lipofectamine 2000 (Thermofisher) was used per 1µg of plasmid DNA. For BRPF2 mutants, 6µg of HA-tagged BRPF2, 6µg FLAG-tagged HBO1 and 3µg of HA-tagged ING5 and 3µg HA-tagged EAF6 are transfected. Plasmids and 18µl of Lipofectamine 2000 were used, in antibiotic free DMEM/FBS media. The mixture is incubated at 37°C under 5% CO2 for 5 to 16 hours. Afterwards, the media is changed to DMEM with FBS and P/S and incubated. 48 hours post-transfection, the media and cells are collected for further analysis.

3.3.4 Immunoprecipitation and Acetylation Assays

Forty-eight hours post transfection, cells were washed twice with PBS and soluble protein extracts were obtained, as previously described by Yan[137], FLAG peptide (Sigma F3290) was used as a means to elute bound proteins from the M2 agarose beads (Millipore Sigma A2220). Both whole cell extracts and affinity purified were prepared for immunoblotting and complex

forming detection using anti-FLAG (Sigma F3165), anti-HA (Biolegend), anti-H3K14ac (EMD Millipore 07-353), anti-H3K23ac (EMD Millipore 07-355), anti-H4 (Abcam ab18253), anti-H4K5ac (Millipore Sigma 07-327), anti-H4K8ac (Millipore Sigma 07-328) antibodies as previously described by Yan *et al.* [137].

Histone acetylation assays were performed based on modified protocol by Yan *et al.* [137]. The purified protein complexes of acetyltransferase were subjected to acetylation reactions using a mixture of purified water, histone or nucleosome substrates, acetyl-CoA, and 5x buffer A. After incubating the reactions at 37 °C for 1 hour, the reactions were stopped, and the samples were used for SDS-PAGE and subsequent immunoblotting. Anti-acetyl lysine antibodies, which were specific to particular histone lysine residues, along with anti-FLAG and HA antibodies, were employed for detection. The next steps are similar to the assay mentioned in the previous chapter.

3.4 Results

Through several international collaborations with several physicians, we have identified nine patients with Intellectual Disability who have undergone Whole Exome Sequencing. Every patient was directed to our laboratory because their WES findings revealed they carry mutations in the *BRPF2* gene. These subjects are all from different families and they are *de novo* mutations. To determine if these variants have the potential to cause the Intellectual Disability phenotype they must be analyzed. All the mutants can be found in Table 3.

Among the *BRPF2* mutations, eight are missense and one is nonsense mutation that lead to C-terminal truncations of the protein (Table 3). Mutation sites are on various regions of BRPF2. The mutants which the ING5- and MEAF6-interacting domain is intact in them are expected to form tetrameric complexes with KAT7, ING5, and MEAF6. Hence, the variations seem to create

unique sets of variants, indicating that these mutations could disrupt BRPF2 functions via various methods.

In understanding that *BRPF2* mutations likely impact function, structure, or stability of *BRPF2*, it is possible that they responsible for causing the clinical features found in the patients, thus we decided to proceed with further analysis of the mutations.

KAT7(HBO1) FLAG-tagged plasmids were generated through PCR amplification of two KAT7 fragments with restriction sites, standard cloning procedures were then followed. The insertion of the full KAT7 gene into the FLAG vector was confirmed by colony PCR, restriction digestion and validated by sequencing. This process was repeated to generate HA-tagged plasmids of the MYST complex, namely, BRPF2, ING5 and EAF6. Mutations in the BRPF2 cDNA were generated via site directed mutagenesis, with primers designed to specifically insert the patient mutation. Mutation insertion was confirmed with sequencing.

The MYST complex plasmids, along with the BRPF2 mutants were used in the following experiments to test the differences in BRPF2. To begin assessing individual differences between the wild type and mutant BRPF1 proteins a transfection of *BRPF2* was performed. For this analysis *BRPF2* plasmids were transfected into HEK293 cells and purified through CO-IP.

The results show that all the BRPF2 variants are able to form the tetrameric complex except for p.Cys327Arg (c.679T>C) and p.Arg504* (c.1509A>T). These two variants fail to form the complex (Fig. 3.5.4). In addition, histone acetylation assays show that these two mentioned variants and p.Arg534Cys (c.1602A>T) indicate less acetylation activity on Histone H4 (Fig. 3.5.5). Other histone acetyl transferase assays are being carried by others in the lab to further investigate the functionality of the variants.

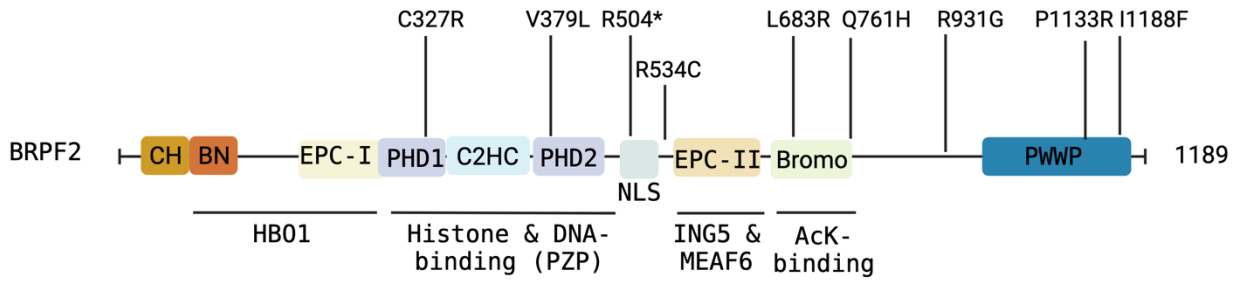
In conclusion, we have identified and analyzed human *BRPF2* mutations. This project sets the stage for further research to firmly establish the direct link of BRPF2 to a new developmental disorder. This chapter thus complements chapter II, which is focused on BRPF1. The links of BRPF1 and BRPF2 in two new developmental disorders also suggest the intriguing possibility that BRPF3 is also mutated in some patients with developmental anomalies. Further research will address this important question.

3.5 Illustrations

Table 3.1 Identification of BRPF2 Mutations in Nine Individuals

Individual	Mutation (GenBank: NM_001394552.1)	Substitution in BRPF2
P1	c.679T>C	p.Cys327Arg
P2	c.1134G>C	p.Val379Leu
P3	c.1509A>T	p.Arg504*
P4	c.1602A>T	p.Arg534Cys
P5	c.2048delT	p.Leu683Arg
P6	c.2283G>C	p.Gln761His
P7	c.2791A>G	p.Arg931Gly
P8	c.3398C>G	p.Pro1133Arg
P9	c.3562A>T	p.Ile1188Phe

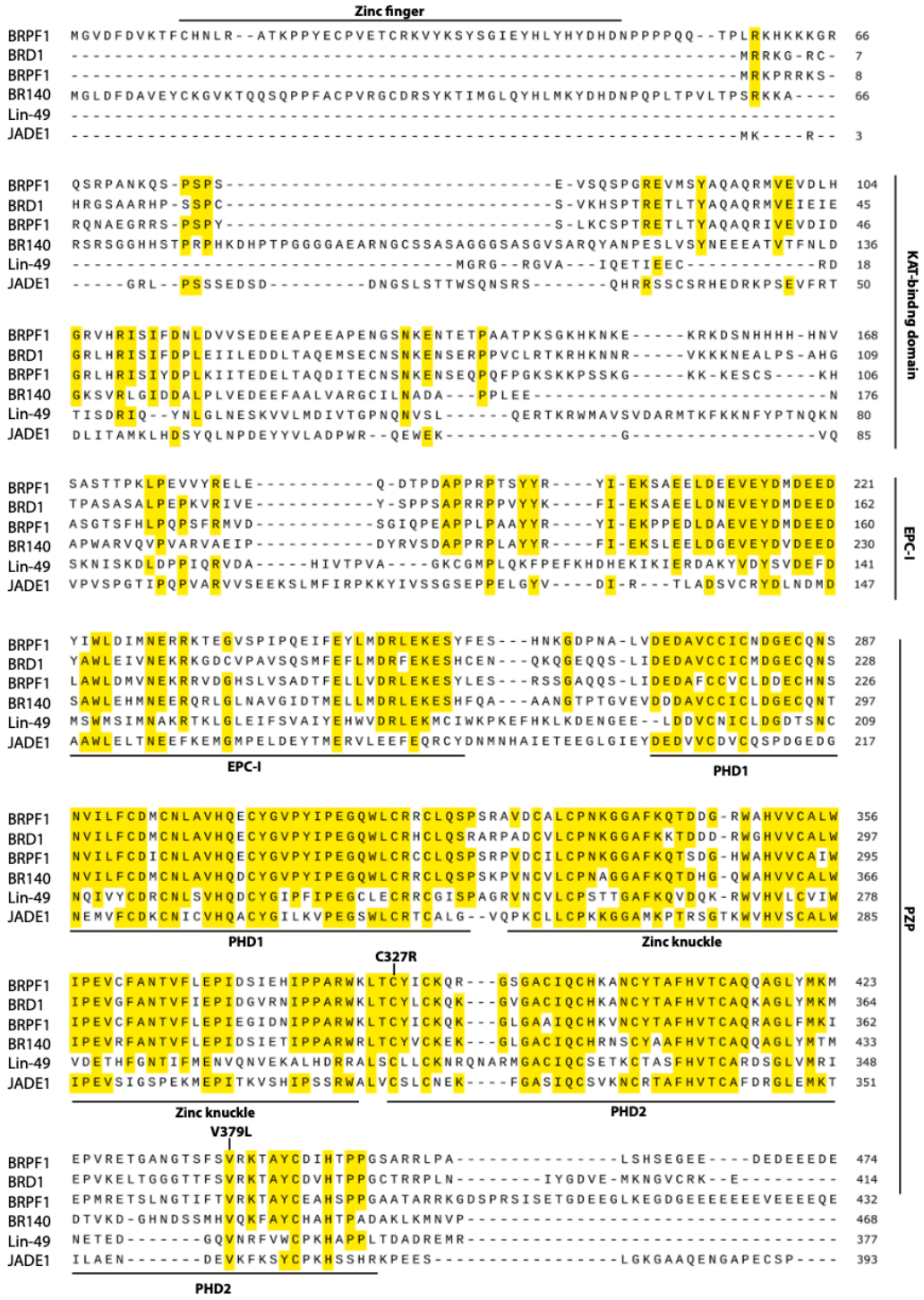
Figure 3.5.1

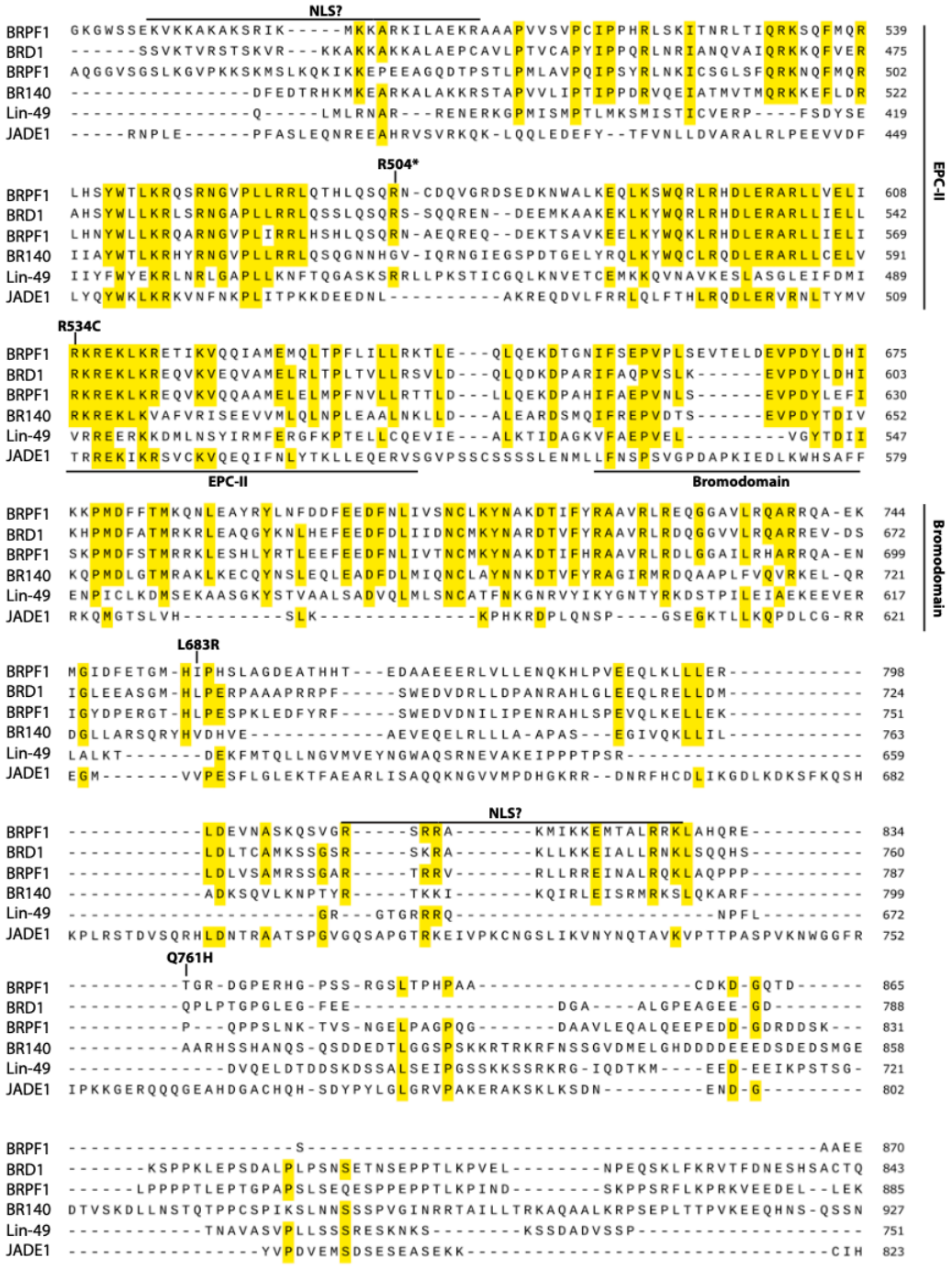


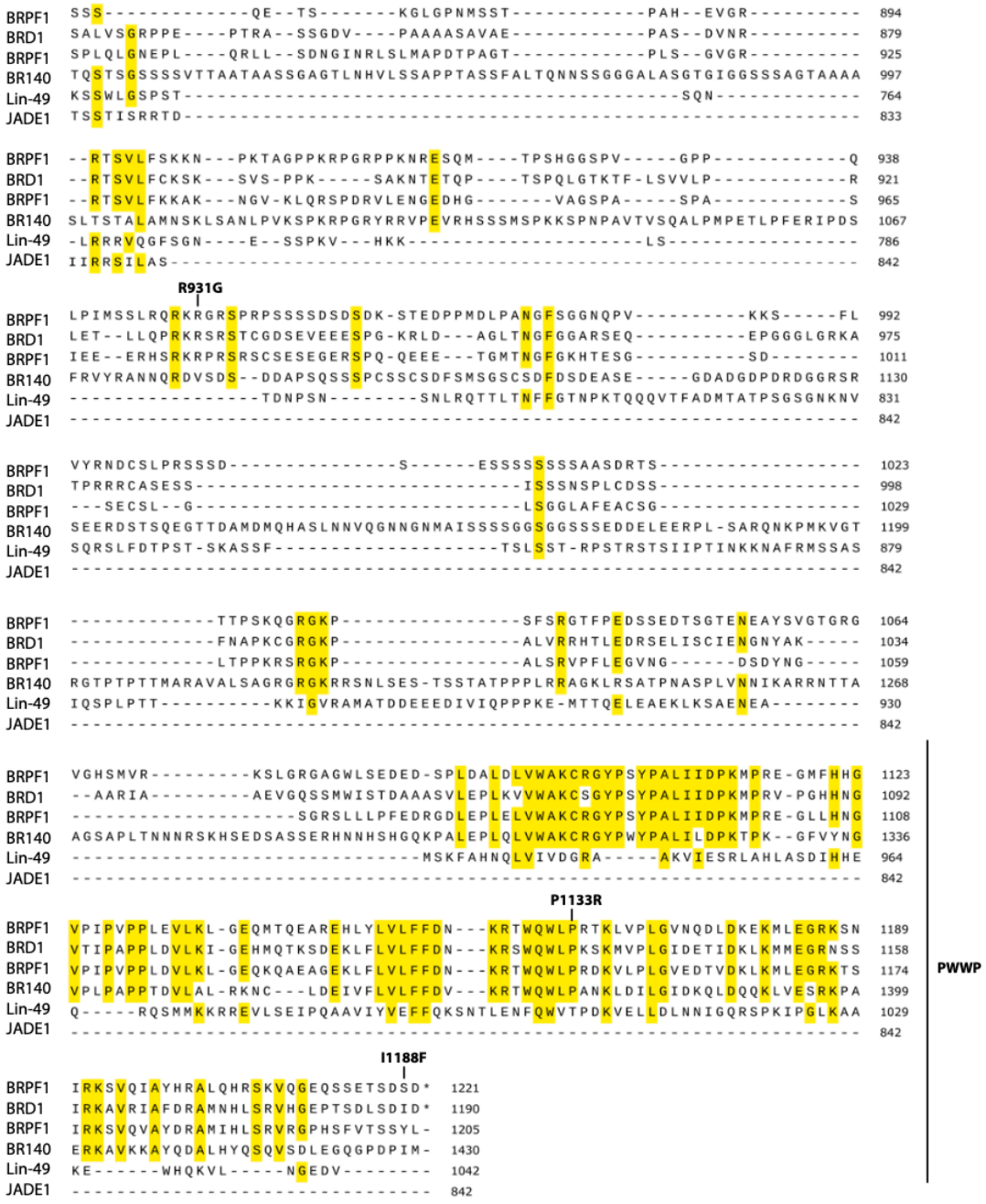
Location of new BRPF2 mutations

The diagrammatic depiction outlines the structure of BRPF2, featuring multiple domain annotations in accordance with identified mutations across various individuals, detailed in the accompanying Table 1. BRPF2 encompasses a consortium of domains integral to its interaction with chromatin components, which includes a CH domain at the N-terminus with potential chromatin binding properties, followed by a BN domain. Notably, the protein houses a PWWP domain, characterized by its specificity towards methylated histone tails, potentially mediating chromatin association. This schematic serves to elucidate the functional territories within BRPF2 and their correlation with the mutations identified in specific individuals. The collective understanding of these domains and mutations provides insights into the broader implications of BRPF2 in chromatin dynamics and gene expression control. Image created with BioRender.com.

Figure 3.5.2


Zinc finger
BRPF1 MGVDFDVKTFCHNLR--ATKPPYECPVETCRKVYKSYSGIEYHLYHYDHDNPPPPQQ--TPLRKHKKKGR 66
BRD1 ----------MRRKG-RC- 7
BRPF1 ----------MRKPRRKS- 8
BR140 MGLDFDAVEYCKGVKTQQSQPPFACPVRGCDRSYKTIMGLQYHLMKYDHDNPQPLTPVLTPSRKKA---- 66
Lin-49 ----------
JADE1 ----------MK----R-- 3
BRPF1 QSRPANKQS-PSPS----------E-VSQSPGREVMSYAQAQRMVEVDLH 104
BRD1 HRGSAARHP-SSPC----------S-VKHSPTRETLTYAQAQRMVEIEIE 45
BRPF1 RQNAEGRRS-PSPY----------S-LKCSPTRETLTYAQAQRIVEVDID 46
BR140 RSRSGGHHSTPRPHKDHPTPGGGGAEARNGCSSASAGGGSASGVSARQYANPESLVSYNEEEATVTFNLD 136
Lin-49 ----------MGRG--RGVA---IQETIEEC----------RD 18
JADE1 -----GRL--PSSSEDSD---------DNGSLSTTWSQNSRS-------QHRRSSCSRHEDRKPSEVFRT 50
KAT-binding domain
BRPF1 GRVHRISIFDNLDVVSEDEEAPEEAPENGSNKENTETPAATPKSGKHKNKE-----KRKDSNHHHH-HNV 168
BRD1 GRLHRISIFDPLEIILEDDLTAQEMSECNSNKENSERPPVCLRTKRHKNNR-----VKKKNEALPS-AHG 109
BRPF1 GRLHRISIYDPLKIITEDELTAQDITECNSNKENSEQPQFPGKSKKPSSKG-----KK-KESCS----KH 106
BR140 GKSVRLGIDDALPLVEDEEFAALVARGCILNADA---PPLEE----------N 176
Lin-49 TISDRIQ--YNLGLNESKVVLMDIVTGPNQNVSL------QERTKRWMAVSVDARMTKFKKNFYPTNQKN 80
JADE1 DLITAMKLHDSYQLNPDEYYVLADPWR--QEWEK----------G----------VQ 85
BRPF1 SASTTPKLPEVVYRELE----------Q-DTPDAPPRPTSYYR----YI-EKSAEELDEEVEYDMDEED 221
BRD1 TPASASALPEPKVRIVE----------Y-SPPSAPRRPPVYYK----FI-EKSAEELDNEVEYDMDEED 162
BRPF1 ASGTSFHLPQPSFRMVD----------SGIQPEAPPLPAAYYR----YI-EKPPEDLDAEVEYDMDEED 160
BR140 APWARVQVPVARVAEIP----------DYRVSDAPPRPLAYYR----FI-EKSLEELDGEVEYDVDEED 230
Lin-49 SKNISKDLDPPIQRVDA-----HIVTPVA----GKCGMPLQKFPEFKHDHEKIKIERDAKYVDYSVDEFD 141
JADE1 VPVSPGTIPQPVARVVSEEKSLMFIRPKKYIVSSGSEPPELGYV----DI-R---TLADSVCRYDLNDMD 147
EPC-I
BRPF1 YIWLDIMNERRKTEGVSPIPQEIFEYLMDRLEKESYFES---HNKGDPNA-LVDEDAVCCICNDGECQNS 287
BRD1 YAWLEIVNEKRKGDCVPAVSQSMFEFLMDRFEKESHCEN---QKQGEQQS-LIDEDAVCCICMDGECQNS 228
BRPF1 LAWLDMVNEKRRVDGHSLVSADTFELLVDRLEKESYLES---RSSGAQQS-LIDEDAFCCVCLDDECHNS 226
BR140 SAWLEHMNEERQRLGLNAVGIDTMELLMDRLEKESHFQA---AANGTPTGVEVDDAVCCICLDGECQNT 297
Lin-49 MSWMSIMNAKRTKLGLEIFSVAIYEHWVDRLEKMCIWKPKEFHKLKDENGEE--LDDVCNICLDGDTSNC 209
JADE1 AAWLELTNEEFKEMGMPELDEYTMERVLEEFEQRCYDNMNHAIETEEGLGIEYDEDVVCDVCQSPDGEDG 217
EPC-I
PHD1
PZP
BRPF1 NVILFCDMCNLAVHQECYGVPYIPEGQWLCRRCLQSPSRAVDCALCPNKGGAFKQTDDG-RWAHVVCALW 356
BRD1 NVILFCDMCNLAVHQECYGVPYIPEGQWLCRHCLQSRARPADCVLCPNKGGAFKKTDDD-RWGHVVCALW 297
BRPF1 NVILFCDICNLAVHQECYGVPYIPEGQWLCRCCLQSPSRPVDCILCPNKGGAFKQTSDG-HWAHVVCAIW 295
BR140 NVILFCDMCNLAVHQDCYGVPYIPEGQWLCRRCLQSPSKPVNCVLCPNAGGAFKQTDHG-QWAHVVCALW 366
Lin-49 NQIVYCDRCNLSVHQDCYGIPFIPEGCLECRRCGISPAGRVNCVLCPSTTGAFKQVDQK-RWVHVLCVIW 278
JADE1 NEMVFCDKCNICVHQACYGILKVPEGSWLCRTCALG--VQPKCLLCPKKGGAMKPTRSGTKWVHVSCALW 285
PHD1
Zinc knuckle
C327R
BRPF1 IPEVCFANTVFLEPIDSIEHIPPARWKLTCYICKQR---GSGACIQCHKANCYTAFHVTCAQQAGLYMKM 423
BRD1 IPEVGFANTVFIEPIDGVRNIPPARWKLTCYLCKQK---GVGACIQCHKANCYTAFHVTCAQKAGLYMKM 364
BRPF1 IPEVCFANTVFLEPIEGIDNIPPARWKLTCYICKQK---GLGAAIQCHKVNCYTAFHVTCAQRAGLFMKI 362
BR140 IPEVRFANTVFLEPIDSIETIPPARWRLTCYVCKEK---GLGACIQCHRNSCYAAFHVTCAQQAGLYMTM 433
Lin-49 VDETHFGNTIFMENVQNVEKALHDRRALSCLLCKNRQNARMGACIQCSETKCTASFHVTCARDSGLVMRI 348
JADE1 IPEVSIGSPEKMEPITKVSHIPSSRWALVCSLCNEK----FGASIQCSVKNCRTAFHVTCAFDRGLEMKT 351
Zinc knuckle
PHD2
V379L
BRPF1 EPVRETGANGTSFSVRKTAYCDIHTPPGSARRLPA----------LSHSEGEE----DEDEEEDE 474
BRD1 EPVKELTGGGTTFSVRKTAYCDVHTPPGCTRRPLN--------IYGDVE-MKNGVCRK--E---------- 414
BRPF1 EPMRETSLNGTIFTVRKTAYCEAHSPPGAATARRKGDSPRSISETGDEEGLKEGDGEEEEEEEVEEEEQE 432
BR140 DTVKD-GHNDSSMHVQKFAYCHAHTPADAKLKMNVP---------- 468
Lin-49 NETED-------GQVNRFVWCPKHAPPLTDADREMR---------- 377
JADE1 ILAEN-------DEVKFKSYCPKHSSHRKPEES----------LGKGAAQENGAPECSP---- 393
PHD2


NLS?
R504*
EPC-II
R534C
EPC-II
Bromodomain
Bromodomain
L683R
NLS?
Q761H

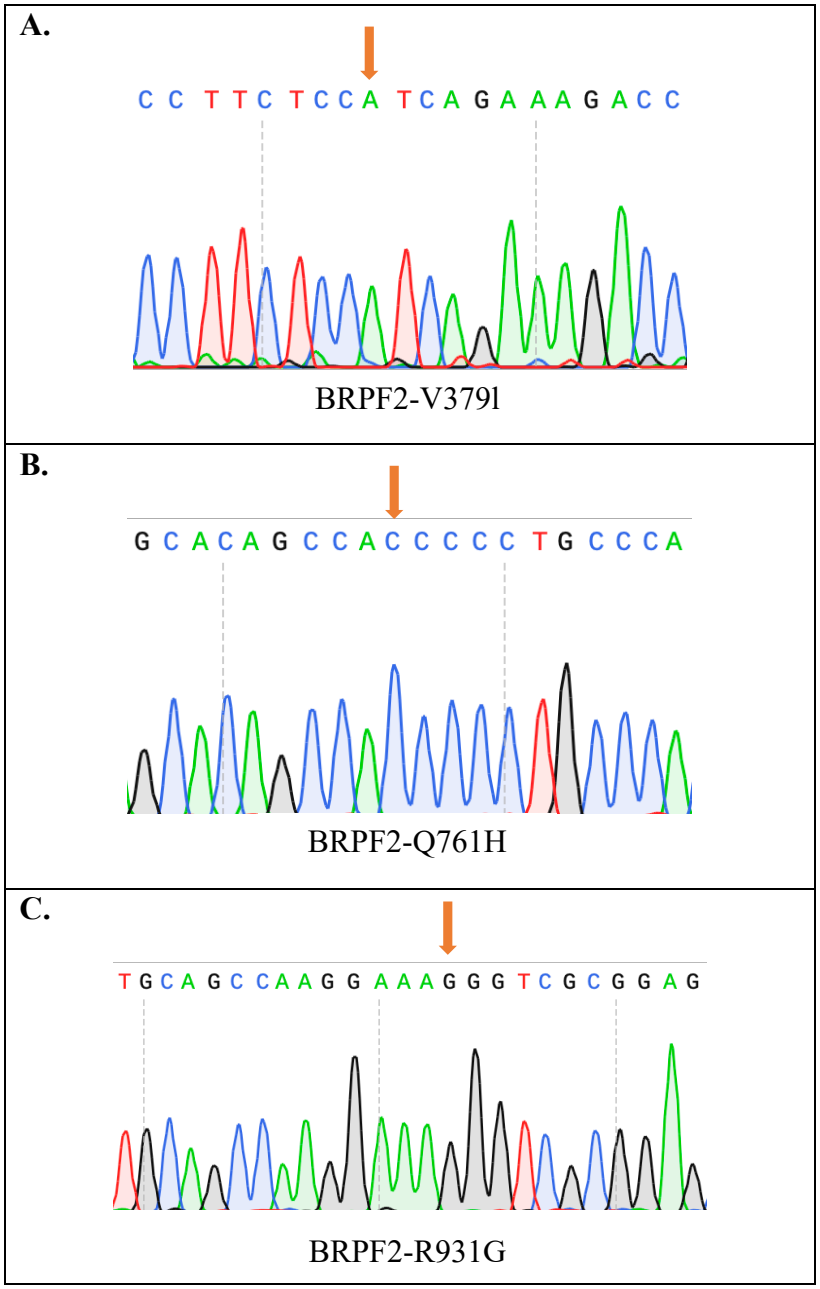


Multiple sequence alignment highlighting conserved regions and mutation sites in BRPF2

This figure presents a multiple sequence alignment of the human Bromodomain and PHD Finger Containing 2 (BRPF2) protein with its paralog Bromodomain containing 1 (BRD1) and

orthologous proteins: LIN-49 from Caenorhabditis elegans, JADE1 (gene Jade-1) and BR140 (also known as BRD7). Conserved amino acid residues across these species are highlighted, indicating regions of structural and functional significance. Key mutation sites are also identified, demonstrating variations that could potentially affect protein function. This comparison allows for the identification of evolutionary conserved domains that are critical for the protein's role in chromatin remodeling and regulation of transcription.

Figure 3.5.3

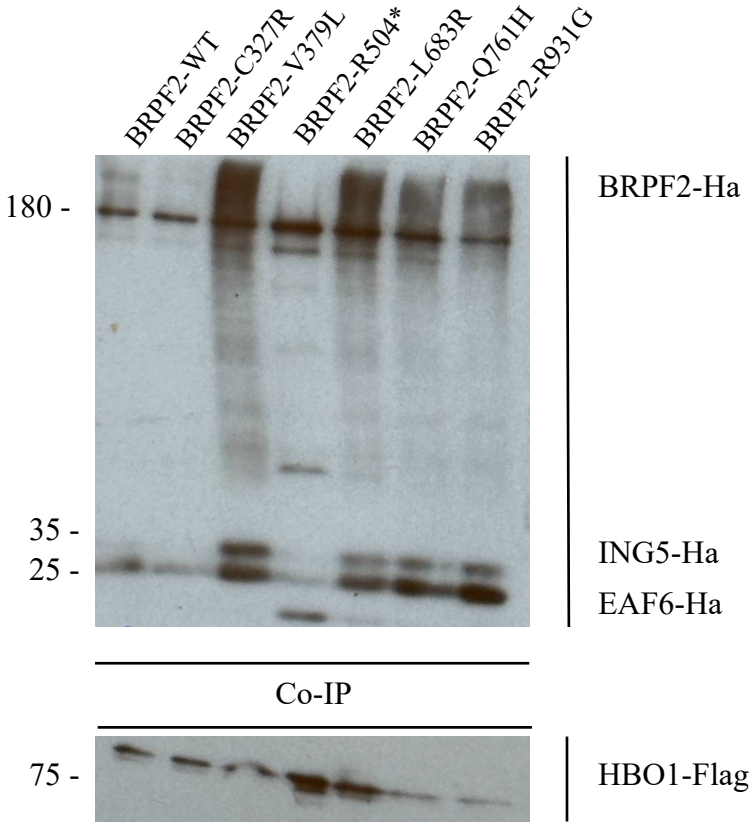


Representative Sanger Sequencing Results

Verification of BRPF2 Mutants by Sanger Sequencing. Panel A, B, and C display the sequencing chromatograms for BRPF2 mutants V379L, Q761H, and R931G respectively. Each panel shows

the specific point mutation in the BRPF2 gene, indicated by an orange arrow, which confirms the successful introduction of the desired mutation. In panel A, the substitution leading to the V379L mutant is evidenced by the peak change from a guanine (G) to a thymine (T). In panel B, the mutation Q761H is identified by the replacement of a cytosine (C) with an adenine (A). Lastly, panel C highlights the R931G mutation through the alteration of a cytosine (C) to a guanine (G). The precise locations of the mutations are demarcated by dashed lines, which correspond to the nucleotide position of the codon change, validating the site-directed mutagenesis process for each of the BRPF2 variants. Chromatograms were analysed with SnapGene.

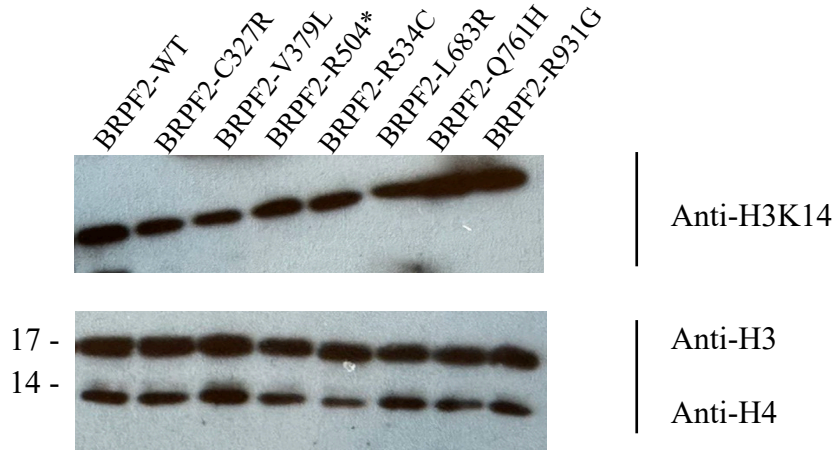
Figure 3.5.4



BRPF2 complex formation with HBO1 (a.k.a. KAT7), ING5 and EAF6

HA-tagged wild type & mutant variants of BRPF1 & ING5 and EAF6 and FLAG-tagged KAT7 underwent transfection in HEK293 cells, followed by IP.

Figure 3.5.5



Histone Acetylation Assays of BRPF2 Variants

Histones acetylation assays were performed using histones utilized as substrates, for affinity purified proteins. Histone acetylation was assessed through the use of specific histone H3 and H4 antibodies. Variations in band intensity across the mutants compared to WT reflect differences in HAT activity attributable to the specific mutations.

Chapter 4 General Conclusion and Disucssion

4.1 General Conclusion

During the course of this project, several key achievements have underscored the advancement in our understanding of the mutagenesis process and its implications for the structure-function relationship of BRPF1 and BRPF2 mutants. Notably, we have gained insights in the following four aspects:

1) Development of a highly effficient mutagenesis method: A robust mutagenesis technique was established, dramatically increasing the efficacy of the mutation introduction process. This methodological enhancement is expected to facilitate future research in genetic modification with improved precision and reliability. Overall, our method has improved the currently repported mutagenesis methods that are being widely used in different labs (see Fig. 2.5.1C). A manuscript on this improved method is currently being prepared for submission (Mousavi, N. & Yang, X.J., unpublished). This is significant, especially when considering the general and wide usage in different other projects in this and many other laboratories.

2) Production of 17 BRPF1 Mutants: 17 BRPF1 mutants were successfully engineered. This achievement represents a considerable expansion of the available mutants for study, laying the groundwork for a deeper exploration of the role of BRPF1 in chromatin modification and interaction with other histone acetyltransferases.

3) Generation of 8 BRPF2 Mutants: The project successfully yielded 8 distinct BRPF2 mutants, thereby enriching the toolkit for dissecting the BRPF2 function and its contribution to chromatin remodeling and reader function.

4) Prelimineray Functional Assays: Complex formation assays confirmed that some mutations affect acetyltransferase complxes. histone acetyltransferase (HAT) assays were executed effectively, providing valuable insights into the enzymatic activity changes due to mutagenesis. These assays are crucial for the elucidation of the structure-function relationship of the histone acetyltransferase complex components. In part due to being carried out during and immediately after the pandemic, one weakness of this study is the relative lack of functional analyses. But this is being improved by others in the lab.

In sum, point 1) above is leading to a manuscript (under preparation). With additional functional assays, points 2-4) above will result in two more manuscripts, with me as the first or co-first author. I have also contributed to a protocol paper [137]. Intellectually, this thesis has shaped my mind for scientific reasoning, which shall help my life and career in the long run.

4.2 General Discussion

Maintaining worldwide epigenetic modification patterns is crucial for sustaining life. The enduring nature of modification patterns and their associated genes is a product of evolutionary preservation. Changes to these genes or the overall epigenome can result in harmful consequences, encompassing conditions like cancer and neurological disorders.

The human genome contains numerous chromatin readers that aid in interpreting distinct states of chromatin modifications. While a portion of these readers has been extensively studied both at a molecular level and in model organisms, only a small fraction has been linked to Mendelian disorders in humans. This study focused on two distinct genes found within histone acetylation complexes: BRPF1 and BRPF2. The objective was to comprehend the consequences of new mutations on protein and epigenomic traits, as well as their effects on patients. The findings presented in this thesis demonstrate that mutations in BRPF1 and BRPF2 lead to a developmental

disorder observed in ten individuals diagnosed with a syndrome characterized by intellectual disability. The developmental delay, intellectual disability, and language impairment align with recent knockout research, indicating the crucial role of mouse Brpf1 in both embryo survival and the development of the forebrain [73, 132, 139] . Certain clinical characteristics, such as intellectual disability and developmental delay, observed in this study are shared with individuals who have mutations in KAT6A or KAT6B [110, 112, 140] .

BRPF1 is conserved across various species, from worms to humans [67] . In C. elegans, a related protein called Lin-49 regulates aspects like neuron asymmetry, hindgut development, and fecundity [141]. Drosophila Br140 is remarkably similar to mammalian BRPF1 and forms a tetrameric complex with Enok, which is equivalent to KAT6A and KAT6B in targeting histone H3K23 acetylation [142] . Enok plays a critical role in neuroblast proliferation in the fly brain and is important for the neuronal wiring of the visual system [82, 143] . While the precise functions of BRPF1 and KAT6A in the nervous system remain unclear, both are essential for maintaining pharyngeal segmental identity and skeletal development in zebrafish [116, 144] . Similarly, both BRPF1 and KAT6B are crucial for mouse brain development [145]. This correlates with the major clinical feature of intellectual disability in individuals with BRPF1 or KAT6B mutations.

It has been found that BRPF1 is vital for mouse hematopoietic stem cells and its inactivation leads to pre-weaning lethality due to bone marrow failure. However, no hematological issues have been detected in individuals with BRPF1 mutations. A similar scenario exists for KAT6A and hematopoietic stem cells. While mice with both alleles inactivated exhibit abnormalities, individuals with one mutated allele do not show indications of cancer predisposition, despite some somatic mutations in BRPF1 being linked to cancer. Nonetheless, further research is needed to explore these intriguing phenomena.

In addition, the intricate involvement of BRPF2 and BRPF3 in regulating HBO1's histone acetyltransferase (HAT) activity underscores the complex molecular choreography essential for histone modification and subsequent gene expression regulation. Their role is especially pivotal in the acetylation of histone H3 and H4, critical processes in the maintenance of chromatin structure and function [146]. A notable aspect of this regulatory mechanism is the influence of these scaffolding proteins on HBO1-mediated HAT activity, particularly concerning H3K14. Deletion of HBO1 in mice models significantly diminishes global H3K14 acetylation, a deficiency that precipitates pronounced developmental anomalies during embryogenesis. This finding highlights the indispensable role of HBO1 and its associated scaffolding proteins in normal embryological development [90].

Parallel to this, the absence of BRPF2 in mice has been observed to lead to a conspicuous reduction in the acetylation of histone H3 lysine 14 at the promoters of genes that are instrumental in the regulation of erythroid development [72]. This outcome signals the importance of BRPF2 in the erythropoietic process, where it possibly maintains the necessary gene expression profile for normal erythroid lineage progression. Nevertheless, the specific mechanism through which BRPF2 orchestrates the recruitment of HBO1 HAT to chromatin, thereby influencing gene transcription, is still an enigma, indicating a gap in our understanding of these proteins' operational dynamics. Recent studies have shed light on the BRPF3 bromodomain, illustrating its ability to identify multiple acetylated lysine residues on histone H4's N-terminal tails. Interestingly, it exhibits a predilection for interactions with H4K5ac and H4K5acK12ac marks , suggesting that these specific histone modifications are of considerable significance in the chromatin remodeling landscape [147]. Given the substantial sequence similarity between the BRPF2 and BRPF3 bromodomains,

an intriguing hypothesis emerges that the BRPF2 bromodomain might also recognize and interact with a comparable set of acetylated histone modifications.

This supposition leads us to speculate that the molecular recognition patterns of BRPF2 might be broader or possibly more nuanced than currently understood. It further raises the question of whether these bromodomains might exhibit functional redundancy or engage in complementary or even divergent roles in chromatin remodeling and gene regulation.

The biological roles and molecular mechanisms of BRPF2 and BRPF3, particularly in the context of histone modification recognition, represent a rich vein for future research. Unraveling these mechanisms will not only enhance our understanding of gene regulation intricacies but also potentially unveil new therapeutic targets for diseases associated with aberrant chromatin remodeling and gene expression.

To encapsulate, our study furnishes significant evidence indicating that the presence of heterozygous mutations in the BRPF1 and BRPF2 gene is responsible for intellectual disabilities and a spectrum of other developmental anomalies. These mutations particularly impinge on the structural and functional rapport between BRPF1 and the specific enzymes KAT6A and KAT6B. This compromised interaction precipitates a marked deficiency in the acetylation process of histone H3 at the lysine 23 position, a critical process for proper chromatin remodeling and gene expression.

The implications of our findings are manifold. First and foremost, they establish a foundational understanding for the phenotypic identification and molecular categorization of developmental disorders associated with BRPF1 mutations. This is a significant stride forward in the realm of medical genetics, where precise diagnosis is pivotal for management and potential therapeutic interventions.

Furthermore, our research underscores the pivotal role of these specific chromatin regulators, BRPF1 and BRPF2 complexes, in the intricate processes underpinning human development and physiological complexity. It opens up new avenues of scientific inquiry into the functions of similar chromatin regulators found in an array of organisms, helping us understand whether these functions are conserved through evolution or if they have diverged to support different physiological processes.

Our research into BRPF1 and BRPF2 mutations dovetails with recent discoveries concerning mutations in KAT6A and KAT6B and KAT7, unearthing an emerging cluster of intellectual disability disorders with a common thread: aberrant histone H3 acetylation. This revelation is not just about expanding the repository of knowledge on genetic disorders; it's about uncovering a novel pathological mechanism underpinning a group of intellectual disabilities. This convergence in etiology highlights an intriguing new aspect of epigenetic regulation in human cognitive development and sets the stage for a potentially unified approach to understanding and perhaps treating these conditions.

In conclusion, we have identified and engineered 25 new *BRPF1* or *BRPF2* mutations derived from patients with neurodevelopment disorders. During the course of this research project, we have also optimized a site-directed mutagenesis method and achieved a high efficiency of 50-100%. Therefore, this project has set up a solid foundation for analysis new *BRPF1* and *BRPF2* mutations for strengthening their links to neurodevelopmental disorders.

References

1. McGinty, R.K. and S. Tan, *Nucleosome structure and function.* Chemical reviews, 2015. **115**(6): p. 2255-2273.
2. Boltsis, I., et al., *Chromatin conformation in development and disease.* Frontiers in Cell and Developmental Biology, 2021. **9**: p. 723859.
3. Ramazi, S. and J. Zahiri, *Post-translational modifications in proteins: resources, tools and prediction methods.* Database, 2021. **2021**.
4. Bidlingmaier, S. and B. Liu, *Identification of posttranslational modification-dependent protein interactions using yeast surface displayed human proteome libraries.* Yeast Surface Display: Methods, Protocols, and Applications, 2015: p. 193-202.
5. Delcuve, G.P., M. Rastegar, and J.R. Davie, *Epigenetic control.* Journal of cellular physiology, 2009. **219**(2): p. 243-250.
6. Waddington, C.H., *The epigenotype.* Endeavour, 1942. **1**: p. 18-20.
7. Felsenfeld, G., *A brief history of epigenetics.* Cold Spring Harbor perspectives in biology, 2014. **6**(1): p. a018200.
8. Bure, I.V., M.V. Nemtsova, and E.B. Kuznetsova, *Histone Modifications and Non-Coding RNAs: Mutual Epigenetic Regulation and Role in Pathogenesis.* International Journal of Molecular Sciences, 2022. **23**(10): p. 5801.
9. Turner, B.M., *Epigenetic responses to environmental change and their evolutionary implications.* Philosophical Transactions of the Royal Society B: Biological Sciences, 2009. **364**(1534): p. 3403-3418.
10. Moore, L.D., T. Le, and G. Fan, *DNA methylation and its basic function.* Neuropsychopharmacology, 2013. **38**(1): p. 23-38.

11. Fitzpatrick, D.R. and C.B. Wilson, *Methylation and demethylation in the regulation of genes, cells, and responses in the immune system.* Clinical immunology, 2003. **109**(1): p. 37-45.

12. Statello, L., et al., *Gene regulation by long non-coding RNAs and its biological functions.* Nature reviews Molecular cell biology, 2021. **22**(2): p. 96-118.

13. Diamantopoulos, M.A., P. Tsiakanikas, and A. Scorilas, *Non-coding RNAs: the riddle of the transcriptome and their perspectives in cancer.* Annals of Translational Medicine, 2018. **6**(12).

14. Jarroux, J., A. Morillon, and M. Pinskaya, *History, discovery, and classification of lncRNAs.* Long non coding RNA biology, 2017: p. 1-46.

15. Catalanotto, C., C. Cogoni, and G. Zardo, *MicroRNA in control of gene expression: an overview of nuclear functions.* International journal of molecular sciences, 2016. **17**(10): p. 1712.

16. Dana, H., et al., *Molecular mechanisms and biological functions of siRNA.* International journal of biomedical science: IJBS, 2017. **13**(2): p. 48.

17. Lee, J.-H., F. Xiong, and W. Li, *Enhancer RNAs in cancer: regulation, mechanisms and therapeutic potential.* RNA biology, 2020. **17**(11): p. 1550-1559.

18. Mandrekar, P., *Epigenetic regulation in alcoholic liver disease.* 2011.

19. Becker, P.B. and J.L. Workman, *Nucleosome remodeling and epigenetics.* Cold Spring Harbor perspectives in biology, 2013. **5**(9): p. a017905.

20. Kadoch, C., *Structure and function of ATP-dependent chromatin remodeling complexes in human cancer*. 2019, American Society of Hematology Washington, DC.

21. Zhou, C.Y., et al., *Mechanisms of ATP-dependent chromatin remodeling motors.* Annual review of biophysics, 2016. **45**: p. 153-181.

22. Tyagi, M., et al., *Chromatin remodelers: We are the drivers!!* Nucleus, 2016. **7**(4): p. 388-404.

23. Bönisch, C. and S.B. Hake, *Histone H2A variants in nucleosomes and chromatin: more or less stable?* Nucleic acids research, 2012. **40**(21): p. 10719-10741.

24. Fan, J., et al., *Metabolic regulation of histone post-translational modifications.* ACS chemical biology, 2015. **10**(1): p. 95-108.

25. Kang, H., et al., *Dynamic regulation of histone modifications and long-range chromosomal interactions during postmitotic transcriptional reactivation.* Genes & Development, 2020. **34**(13-14): p. 913-930.

26. Grimes, M., et al., *Integration of protein phosphorylation, acetylation, and methylation data sets to outline lung cancer signaling networks.* Science signaling, 2018. **11**(531): p. eaaq1087.

27. Paik, W.K., D.C. Paik, and S. Kim, *Historical review: the field of protein methylation.* Trends in biochemical sciences, 2007. **32**(3): p. 146-152.

28. Black, J.C., C. Van Rechem, and J.R. Whetstine, *Histone lysine methylation dynamics: establishment, regulation, and biological impact.* Molecular cell, 2012. **48**(4): p. 491-507.

29. Miller, J.L. and P.A. Grant, *The role of DNA methylation and histone modifications in transcriptional regulation in humans.* Epigenetics: development and disease, 2012: p. 289-317.

30. Jambhekar, A., A. Dhall, and Y. Shi, *Roles and regulation of histone methylation in animal development.* Nature reviews Molecular cell biology, 2019. **20**(10): p. 625-641.

31. Jin, J. and T. Pawson, *Modular evolution of phosphorylation-based signalling systems.* Philosophical Transactions of the Royal Society B: Biological Sciences, 2012. **367**(1602): p. 2540-2555.

32. Zhou, Z., et al., *Therapy of Infectious Diseases Using Epigenetic Approaches.* Epigenetics in Human Disease, 2018: p. 689-715.

33. Oberdoerffer, P. and K.M. Miller. *Histone H2A variants: Diversifying chromatin to ensure genome integrity*. in *Seminars in cell & developmental biology*. 2023. Elsevier.

34. Jo, C., et al., *Histone acylation marks respond to metabolic perturbations and enable cellular adaptation.* Experimental & molecular medicine, 2020. **52**(12): p. 2005-2019.

35. Sabari, B.R., et al., *Metabolic regulation of gene expression through histone acylations.* Nature reviews Molecular cell biology, 2017. **18**(2): p. 90-101.

36. Nitsch, S., L. Zorro Shahidian, and R. Schneider, *Histone acylations and chromatin dynamics: concepts, challenges, and links to metabolism.* EMBO reports, 2021. **22**(7): p. e52774.

37. Cao, J., et al., *HDAC11 regulates type I interferon signaling through defatty-acylation of SHMT2.* Proceedings of the National Academy of Sciences, 2019. **116**(12): p. 5487-5492.

38. Grant, P.A., *A tale of histone modifications.* Genome biology, 2001. **2**: p. 1-6.

39. Galdieri, L., et al., *Protein acetylation and acetyl coenzyme a metabolism in budding yeast.* Eukaryotic cell, 2014. **13**(12): p. 1472-1483.

40. Dong, X. and Z. Weng, *The correlation between histone modifications and gene expression.* Epigenomics, 2013. **5**(2): p. 113-116.

41. Yang, X.J., *The diverse superfamily of lysine acetyltransferases and their roles in leukemia and other diseases.* Nucleic acids research, 2004. **32**(3): p. 959-976.

42. Park, S.-Y. and J.-S. Kim, *A short guide to histone deacetylases including recent progress on class II enzymes.* Experimental & molecular medicine, 2020. **52**(2): p. 204-212.

43. Seto, E. and M. Yoshida, *Erasers of histone acetylation: the histone deacetylase enzymes.* Cold Spring Harbor perspectives in biology, 2014. **6**(4): p. a018713.

44. Moosavi, A. and A. Motevalizadeh Ardekani, *Role of epigenetics in biology and human diseases.* Iranian biomedical journal, 2016. **20**(5): p. 246-258.

45. Schreiber, S.L. and B.E. Bernstein, *Signaling network model of chromatin.* Cell, 2002. **111**(6): p. 771-778.

46. Bray, F., et al., *The ever-increasing importance of cancer as a leading cause of premature death worldwide.* Cancer, 2021. **127**(16): p. 3029-3030.

47. Sadikovic, B., et al., *Cause and consequences of genetic and epigenetic alterations in human cancer.* Current genomics, 2008. **9**(6): p. 394-408.

48. Sharma, S., T.K. Kelly, and P.A. Jones, *Epigenetics in cancer.* Carcinogenesis, 2010. **31**(1): p. 27-36.

49. Ng, J.M.-K. and J. Yu, *Promoter hypermethylation of tumour suppressor genes as potential biomarkers in colorectal cancer.* International journal of molecular sciences, 2015. **16**(2): p. 2472-2496.

50. Riggs, A.D. and P.A. Jones, *5-methylcytosine, gene regulation, and cancer.* Advances in cancer research, 1983. **40**: p. 1-30.

51. Jones, P.A. and S.B. Baylin, *The fundamental role of epigenetic events in cancer.* Nature reviews genetics, 2002. **3**(6): p. 415-428.

52. Audia, J.E. and R.M. Campbell, *Histone modifications and cancer.* Cold Spring Harbor perspectives in biology, 2016. **8**(4): p. a019521.

53. Einav Nili, G.-Y., et al., *Cancer epigenetics: modifications, screening, and therapy.* Annu. Rev. Med., 2008. **59**: p. 267-280.

54. Tecalco-Cruz, A.C., et al., *Epigenetic basis of Alzheimer disease.* World Journal of Biological Chemistry, 2020. **11**(2): p. 62.

55. Nikolac Perkovic, M., et al., *Epigenetics of Alzheimer's disease.* Biomolecules, 2021. **11**(2): p. 195.

56. Williams, S.R., et al., *Haploinsufficiency of HDAC4 causes brachydactyly mental retardation syndrome, with brachydactyly type E, developmental delays, and behavioral problems.* The American Journal of Human Genetics, 2010. **87**(2): p. 219-228.

57. Smith, B.C. and J.M. Denu, *Chemical mechanisms of histone lysine and arginine modifications.* Biochimica et Biophysica Acta (BBA)-Gene Regulatory Mechanisms, 2009. **1789**(1): p. 45-57.

58. Sterner, D.E. and S.L. Berger, *Acetylation of histones and transcription-related factors.* Microbiology and molecular biology reviews, 2000. **64**(2): p. 435-459.

59. Annunziato, A.T. and J.C. Hansen, *Role of histone acetylation in the assembly and modulation of chromatin structures.* Gene Expression The Journal of Liver Research, 2001. **9**(1-2): p. 37-61.

60. Musselman, C.A., et al., *Perceiving the epigenetic landscape through histone readers.* Nature structural & molecular biology, 2012. **19**(12): p. 1218-1227.

61. Yan, K., et al., *Mutations in the chromatin regulator gene BRPF1 cause syndromic intellectual disability and deficient histone acetylation.* The American Journal of Human Genetics, 2017. **100**(1): p. 91-104.

62. Wahab, S., et al., *Exploring the histone acetylation cycle in the protozoan model Tetrahymena thermophila.* Frontiers in Cell and Developmental Biology, 2020. **8**: p. 509.

63. Marmorstein, R. and M.-M. Zhou, *Writers and readers of histone acetylation: structure, mechanism, and inhibition.* Cold Spring Harbor perspectives in biology, 2014. **6**(7): p. a018762.

64. Kim, G.-W. and X.-J. Yang, *Comprehensive lysine acetylomes emerging from bacteria to humans.* Trends in biochemical sciences, 2011. **36**(4): p. 211-220.

65. Berndsen, C.E. and J.M. Denu, *Catalysis and substrate selection by histone/protein lysine acetyltransferases.* Current opinion in structural biology, 2008. **18**(6): p. 682-689.

66. Utley, R. and J. Cote, *The MYST family of histone acetyltransferases.* Protein Complexes That Modify Chromatin, 2003: p. 203-236.

67. Yang, X.-J., *MOZ and MORF acetyltransferases: Molecular interaction, animal development and human disease.* Biochimica et Biophysica Acta (BBA)-Molecular Cell Research, 2015. **1853**(8): p. 1818-1826.

68. Cánovas, D., et al., *The histone acetyltransferase GcnE (GCN5) plays a central role in the regulation of Aspergillus asexual development.* Genetics, 2014. **197**(4): p. 1175-1189.

69. Joanna, F., et al., *Histone deacetylase inhibition and the regulation of cell growth with particular reference to liver pathobiology.* Journal of Cellular and Molecular Medicine, 2009. **13**(9b): p. 2990-3005.

70. Hasan, S., et al., *Transcription coactivator p300 binds PCNA and may have a role in DNA repair synthesis.* Nature, 2001. **410**(6826): p. 387-391.

71. Iizuka, M., et al., *Regulation of replication licensing by acetyltransferase Hbo1.* Molecular and cellular biology, 2006. **26**(3): p. 1098-1108.

72. Mishima, Y., et al., *The Hbo1-Brd1/Brpf2 complex is responsible for global acetylation of H3K14 and required for fetal liver erythropoiesis.* Blood, The Journal of the American Society of Hematology, 2011. **118**(9): p. 2443-2453.

73. You, L., et al., *The chromatin regulator Brpf1 regulates embryo development and cell proliferation.* Journal of Biological Chemistry, 2015. **290**(18): p. 11349-11364.

74. Klein, B.J., et al., *Crosstalk between epigenetic readers regulates the MOZ/MORF HAT complexes.* Epigenetics, 2014. **9**(2): p. 186-193.

75. Bahari-Javan, S., F. Sananbenesi, and A. Fischer, *Histone-acetylation: a link between Alzheimer's disease and post-traumatic stress disorder?* Frontiers in neuroscience, 2014. **8**: p. 160.

76. Feng, Y., J. Jankovic, and Y.-C. Wu, *Epigenetic mechanisms in Parkinson's disease.* Journal of the neurological sciences, 2015. **349**(1-2): p. 3-9.

77. Park, H.-S., et al., *Epigenetic targeting of histone deacetylases in diagnostics and treatment of depression.* International Journal of Molecular Sciences, 2021. **22**(10): p. 5398.

78. Borrow, J., et al., *The translocation t (8; 16)(p11; p13) of acute myeloid leukaemia fuses a putative acetyltransferase to the CREB–binding protein.* Nature genetics, 1996. **14**(1): p. 33-41.

79. Champagne, N., et al., *Identification of a human histone acetyltransferase related to monocytic leukemia zinc finger protein.* Journal of Biological Chemistry, 1999. **274**(40): p. 28528-28536.

80. Ullah, M., et al., *Molecular architecture of quartet MOZ/MORF histone acetyltransferase complexes.* Molecular and cellular biology, 2008. **28**(22): p. 6828-6843.

81. Carlson, S. and K.C. Glass, *The MOZ histone acetyltransferase in epigenetic signaling and disease.* Journal of cellular physiology, 2014. **229**(11): p. 1571-1574.

82. Scott, E.K., T. Lee, and L. Luo, *enok encodes a Drosophila putative histone acetyltransferase required for mushroom body neuroblast proliferation.* Current Biology, 2001. **11**(2): p. 99-104.

83. Huang, F., et al., *Histone acetyltransferase Enok regulates oocyte polarization by promoting expression of the actin nucleation factor spire.* Genes & Development, 2014. **28**(24): p. 2750-2763.

84. Klein, B.J., et al., *Histone H3K23-specific acetylation by MORF is coupled to H3K14 acylation.* Nature communications, 2019. **10**(1): p. 4724.

85. Ali, M., et al., *Tandem PHD fingers of MORF/MOZ acetyltransferases display selectivity for acetylated histone H3 and are required for the association with chromatin.* Journal of molecular biology, 2012. **424**(5): p. 328-338.

86. Iizuka, M. and B. Stillman, *Histone acetyltransferase HBO1 interacts with the ORC1 subunit of the human initiator protein.* Journal of Biological Chemistry, 1999. **274**(33): p. 23027-23034.

87. Avvakumov, N., et al., *Conserved molecular interactions within the HBO1 acetyltransferase complexes regulate cell proliferation.* Molecular and cellular biology, 2012. **32**(3): p. 689-703.

88. Saksouk, N., et al., *HBO1 HAT complexes target chromatin throughout gene coding regions via multiple PHD finger interactions with histone H3 tail.* Molecular cell, 2009. **33**(2): p. 257-265.

89. Xiao, Y., et al., *HBO1 is a versatile histone acyltransferase critical for promoter histone acylations.* Nucleic acids research, 2021. **49**(14): p. 8037-8059.

90. Kueh, A.J., et al., *HBO1 is required for H3K14 acetylation and normal transcriptional activity during embryonic development.* Molecular and cellular biology, 2011. **31**(4): p. 845-860.

91. Kamine, J., et al., *Identification of a cellular protein that specifically interacts with the essential cysteine region of the HIV-1 Tat transactivator.* Virology, 1996. **216**(2): p. 357-366.

92. Tominaga, K., et al., *Tip60/KAT5 Histone Acetyltransferase Is Required for Maintenance and Neurogenesis of Embryonic Neural Stem Cells.* International journal of molecular sciences, 2023. **24**(3): p. 2113.

93. Humbert, J., et al., *De novo KAT5 variants cause a syndrome with recognizable facial dysmorphisms, cerebellar atrophy, sleep disturbance, and epilepsy.* The American Journal of Human Genetics, 2020. **107**(3): p. 564-574.

94. Tang, Y., et al., *Tip60-dependent acetylation of p53 modulates the decision between cell-cycle arrest and apoptosis.* Molecular cell, 2006. **24**(6): p. 827-839.

95. Hilfiker, A., et al., *mof, a putative acetyl transferase gene related to the Tip60 and MOZ human genes and to the SAS genes of yeast, is required for dosage compensation in Drosophila.* The EMBO journal, 1997. **16**(8): p. 2054-2060.

96. Taipale, M., et al., *hMOF histone acetyltransferase is required for histone H4 lysine 16 acetylation in mammalian cells.* Molecular and cellular biology, 2005. **25**(15): p. 6798-6810.

97. Cai, Y., et al., *Subunit composition and substrate specificity of a MOF-containing histone acetyltransferase distinct from the male-specific lethal (MSL) complex.* Journal of Biological Chemistry, 2010. **285**(7): p. 4268-4272.

98. Li, X., et al., *Two mammalian MOF complexes regulate transcription activation by distinct mechanisms.* Molecular cell, 2009. **36**(2): p. 290-301.

99. Sykes, S.M., et al., *Acetylation of the p53 DNA-binding domain regulates apoptosis induction.* Molecular cell, 2006. **24**(6): p. 841-851.

100. Thompson, K.A., et al., *BR140, a novel zinc-finger protein with homology to the TAF250 subunit of TFIID.* Biochemical and biophysical research communications, 1994. **198**(3): p. 1143-1152.

101. Sanchez, R., J. Meslamani, and M.-M. Zhou, *The bromodomain: from epigenome reader to druggable target.* Biochimica et Biophysica Acta (BBA)-Gene Regulatory Mechanisms, 2014. **1839**(8): p. 676-685.

102. Vezzoli, A., et al., *Molecular basis of histone H3K36me3 recognition by the PWWP domain of Brpf1.* Nature structural & molecular biology, 2010. **17**(5): p. 617-619.

103. Wang, Y., et al., *Regulation of Set9-mediated H4K20 methylation by a PWWP domain protein.* Molecular cell, 2009. **33**(4): p. 428-437.

104. Bjarkam, C.R., et al., *Further immunohistochemical characterization of BRD1 a new susceptibility gene for schizophrenia and bipolar affective disorder.* Brain Structure and Function, 2009. **214**: p. 37-47.

105. Zhou, M.I., et al., *The von Hippel-Lindau tumor suppressor stabilizes novel plant homeodomain protein Jade-1.* Journal of Biological Chemistry, 2002. **277**(42): p. 39887-39898.

106. Panagopoulos, I., et al., *Fusion of the MORF and CBP genes in acute myeloid leukemia with the t (10; 16)(q22; p13).* Human molecular genetics, 2001. **10**(4): p. 395-404.

107. Moore, S.D., et al., *Uterine leiomyomata with t (10; 17) disrupt the histone acetyltransferase MORF.* Cancer research, 2004. **64**(16): p. 5570-5577.

108. Lynch, H., et al., *Can unknown predisposition in familial breast cancer be family-specific?* The breast journal, 2013. **19**(5): p. 520-528.

109. Katsumoto, T., et al., *MOZ is essential for maintenance of hematopoietic stem cells.* Genes & development, 2006. **20**(10): p. 1321-1330.

110. Kraft, M., et al., *Disruption of the histone acetyltransferase MYST4 leads to a Noonan syndrome–like phenotype and hyperactivated MAPK signaling in humans and mice.* The Journal of clinical investigation, 2011. **121**(9).

111. Szakszon, K., et al., *De novo mutations of the gene encoding the histone acetyltransferase KAT6B in two patients with Say-Barber/Biesecker/Young-Simpson syndrome.* American journal of medical genetics Part A, 2013. **161**(4): p. 884-888.

112. Simpson, M.A., et al., *De novo mutations of the gene encoding the histone acetyltransferase KAT6B cause Genitopatellar syndrome.* The American Journal of Human Genetics, 2012. **90**(2): p. 290-294.

113. Yu, H.C., et al., *An individual with blepharophimosis–ptosis–epicanthus inversus syndrome (BPES) and additional features expands the phenotype associated with mutations in KAT6B.* American Journal of Medical Genetics Part A, 2014. **164**(4): p. 950-957.

114. Tham, E., et al., *Dominant mutations in KAT6A cause intellectual disability with recognizable syndromic features.* The American Journal of Human Genetics, 2015. **96**(3): p. 507-513.

115. Yang, X. and M. Ullah, *MOZ and MORF, two large MYSTic HATs in normal and cancer stem cells.* Oncogene, 2007. **26**(37): p. 5408-5419.

116. Laue, K., et al., *The multidomain protein Brpf1 binds histones and is required for Hox gene expression and segmental identity.* 2008.

117. Yan, K., et al., *The chromatin regulator BRPF3 preferentially activates the HBO1 acetyltransferase but is dispensable for mouse development and survival.* Journal of Biological Chemistry, 2016. **291**(6): p. 2647-2663.

118. Kim, H.-J. and S.-C. Bae, *Histone deacetylase inhibitors: molecular mechanisms of action and clinical trials as anti-cancer drugs.* American journal of translational research, 2011. **3**(2): p. 166.

119. Ahuja, N., A.R. Sharma, and S.B. Baylin, *Epigenetic therapeutics: a new weapon in the war against cancer.* Annual review of medicine, 2016. **67**: p. 73-89.

120. Yan, K., et al., *Deficient histone H3 propionylation by BRPF1-KAT6 complexes in neurodevelopmental disorders and cancer.* Science advances, 2020. **6**(4): p. eaax0021.

121. Yoshida, M., et al., *Potent and specific inhibition of mammalian histone deacetylase both in vivo and in vitro by trichostatin A.* Journal of Biological Chemistry, 1990. **265**(28): p. 17174-17179.

122. Duvic, M., et al., *Phase 2 trial of oral vorinostat (suberoylanilide hydroxamic acid, SAHA) for refractory cutaneous T-cell lymphoma (CTCL).* Blood, 2007. **109**(1): p. 31-39.

123. C Mwakwari, S., et al., *Macrocyclic histone deacetylase inhibitors.* Current topics in medicinal chemistry, 2010. **10**(14): p. 1423-1440.

124. Dokmanovic, M. and P.A. Marks, *Prospects: histone deacetylase inhibitors.* Journal of cellular biochemistry, 2005. **96**(2): p. 293-304.

125. Rasheed, W.K., R.W. Johnstone, and H.M. Prince, *Histone deacetylase inhibitors in cancer therapy.* Expert opinion on investigational drugs, 2007. **16**(5): p. 659-678.

126. Huang, F., S.M. Abmayr, and J.L. Workman, *Regulation of KAT6 acetyltransferases and their roles in cell cycle progression, stem cell maintenance, and human disease.* Molecular and cellular biology, 2016. **36**(14): p. 1900-1907.

127. Sheikh, B., et al., *MOZ (MYST3, KAT6A) inhibits senescence via the INK4A-ARF pathway.* Oncogene, 2015. **34**(47): p. 5807-5820.

128. Sheikh, B.N., et al., *MOZ regulates B-cell progenitors and, consequently, Moz haploinsufficiency dramatically retards MYC-induced lymphoma development.* Blood, The Journal of the American Society of Hematology, 2015. **125**(12): p. 1910-1921.

129. Baell, J.B., et al., *Inhibitors of histone acetyltransferases KAT6A/B induce senescence and arrest tumour growth.* Nature, 2018. **560**(7717): p. 253-257.

130. Meier, J.C., et al., *Selective targeting of bromodomains of the bromodomain-PHD fingers family impairs osteoclast differentiation.* ACS chemical biology, 2017. **12**(10): p. 2619-2630.

131. Cheng, C.L.-H., et al., *Bromodomain-containing protein BRPF1 is a therapeutic target for liver cancer.* Communications Biology, 2021. **4**(1): p. 888.

132. You, L., et al., *Deficiency of the chromatin regulator BRPF1 causes abnormal brain development.* Journal of Biological Chemistry, 2015. **290**(11): p. 7114-7129.

133. Zhang, C., et al., *BRPF1 bridges H3K4me3 and H3K23ac in human embryonic stem cells and is essential to pluripotency.* Iscience, 2023. **26**(2).

134. Su, Y., et al., *Brpf1 haploinsufficiency impairs dendritic arborization and spine formation, leading to cognitive deficits.* Frontiers in Cellular Neuroscience, 2019. **13**: p. 249.

135. Yan, K., et al., *DDD Study. CAUSES Study Mutations in the chromatin regulator gene BRPF1 cause syndromic intellectual disability and deficient histone acetylation.* Am. J. Hum. Genet, 2017. **100**: p. 91-104.

136. Liu, H. and J.H. Naismith, *An efficient one-step site-directed deletion, insertion, single and multiple-site plasmid mutagenesis protocol.* BMC Biotechnology, 2008. **8**(1): p. 91.

137. Yan, K., N. Mousavi, and X.J. Yang, *Analysis of Lysine Acetylation and Acetylation-like Acylation In Vitro and In Vivo.* Current Protocols, 2023. **3**(5): p. e738.

138. Liu, H. and J.H. Naismith, *An efficient one-step site-directed deletion, insertion, single and multiple-site plasmid mutagenesis protocol.* BMC biotechnology, 2008. **8**(1): p. 1-10.

139. You, L., et al., *The lysine acetyltransferase activator Brpf1 governs dentate gyrus development through neural stem cells and progenitors.* PLoS genetics, 2015. **11**(3): p. e1005034.

140. Clayton-Smith, J., et al., *Whole-exome-sequencing identifies mutations in histone acetyltransferase gene KAT6B in individuals with the Say-Barber-Biesecker variant of Ohdo syndrome.* The American Journal of Human Genetics, 2011. **89**(5): p. 675-681.

141. O'Meara, M.M., F. Zhang, and O. Hobert, *Maintenance of neuronal laterality in Caenorhabditis elegans through MYST histone acetyltransferase complex components LSY-12, LSY-13 and LIN-49.* Genetics, 2010. **186**(4): p. 1497-1502.

142. Huang, F., et al., *The Enok acetyltransferase complex interacts with Elg1 and negatively regulates PCNA unloading to promote the G1/S transition.* Genes & development, 2016. **30**(10): p. 1198-1210.

143. Berger, J., et al., *Systematic identification of genes that regulate neuronal wiring in the Drosophila visual system.* PLoS genetics, 2008. **4**(5): p. e1000085.

144. Miller, C.T., L. Maves, and C.B. Kimmel, *moz regulates Hox expression and pharyngeal segmental identity in zebrafish.* 2004.

145. Thomas, T., et al., *Querkopf, a MYST family histone acetyltransferase, is required for normal cerebral cortex development.* Development, 2000. **127**(12): p. 2537-2548.

146. Tao, Y., et al., *Structural and mechanistic insights into regulation of HBO1 histone acetyltransferase activity by BRPF2.* Nucleic acids research, 2017. **45**(10): p. 5707-5719.

147. Barman, S., et al., *Insights into the molecular mechanisms of histone code recognition by the BRPF3 bromodomain.* Chemistry–An Asian Journal, 2021. **16**(21): p. 3404-3412.