CHARACTERIZATION OF NOVEL SWI/SNF CHROMATIN REMODELING COMPLEX (GBAF) IN HEALTH AND DISEASE

by

Aktan Alpsoy

A Dissertation

Submitted to the Faculty of Purdue University

In Partial Fulfillment of the Requirements for the degree of

Doctor of Philosophy

Department of Medicinal Chemistry and Molecular Pharmacology

West Lafayette, Indiana

May 2020

THE PURDUE UNIVERSITY GRADUATE SCHOOL
STATEMENT OF COMMITTEE APPROVAL

Dr. Emily C. Dykhuizen, Chair

Department of Medicinal Chemistry and Molecular Pharmacology

Dr. Andrea Kasinski

Department of Biological Sciences

Dr. Elizabeth Tran

Department of Biochemistry

Dr. Michael K. Wendt

Department of Medicinal Chemistry and Molecular Pharmacology

Approved by:

Dr. Andy Hudmon

To my precious parents Ipek and Necat and my sisters Ayla and Aysen,

&

biricik Ekin, Kerem and Ilke,

ACKNOWLEDGMENTS

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

In eukaryotic systems, the genetic material of the cell –DNA– is packed into a protein-dense structure called chromatin. Chromatin structure is critical for preservation of the genetic material as well as coordination of vital processes such as DNA replication, transcription and DNA damage repair. The fundamental repeating unit of chromatin is nucleosome which is composed of an octamer of small alkaline proteins called histones and the DNA wrapped around this octamer. The nucleosomes are then packed into higher-order structures leading to formation of 3D chromatin architecture. The chromatin is a dynamic structure; the spacing between nucleosomes, or the folding of the larger chromatin segments is subjected to alterations during embryonic development, tissue specifications or *simply during any event that require gene expression changes*. Failure in proper regulation of chromatin structure has been associated with embryonic defects and disease such as cancer.

This work has focused on a class of ATP-dependent chromatin remodeling complexes known as switch/sucrose-non-fermentable (SWI/SNF) or BRG-associated factors (BAF) complex. This family of complexes act on chromatin and alter its physical structure by mobilizing histones or nucleosome particles through the activity of its ATPase –BRG1 or BRM, enabling more accessible DNA for the other factors such as transcription factors to localize and recruit transcription machinery. In particular, we discovered and biochemically defined a novel version of this family of chromatin complexes that we named as GLTSCR1/1L-BAF (GBAF). GLTSCR1 and GLTSCR1L are two uncharacterized paralogous proteins that have been identified as BRG1-interacting proteins. Biochemically surveying the essence of this interaction, we realized that these proteins incorporates into a previously unknown SWI/SNF family complex that lacks well-characterized SWI/SNF subunits such as ARID1/2, BAF170, BAF47; instead, uniquely comprise GLTSCR1/1L and bromodomain-containing protein BRD9. Focusing on the GLTSCR1 subunit, we observed that its absence is well-tolerated by many different cell types except slight growth retardation by prostate cancer cells. Expanding the cohort of prostate cancer cells, we realized that not the paralogous subunits GLTSCR1 or GLTSCR1L but unique and non-redundant subunit BRD9 is the major GBAF-dependence in prostate cancer cells. We observed that especially the androgen-receptor positive cell lines have severe growth defects upon *BRD9* knockdown or inhibition. *In vivo,* we showed that xenografts with *BRD9* knockdown prostate cancer cells

12

(LNCaP) have smaller tumor size. We demonstrated that BRD9 inhibition can block the expression of androgen-receptor targets. Similarly, *BRD9* knockdown and treatment with antiandrogen drug (enzalutamide) has overlapping transcriptional effects. Mechanistically, we showed that BRD9 interacts with AR and it colocalizes with AR in subset of AR -binding sites. Surprisingly, we realized that BRD9 depletion has similar transcriptional and phenotypic effects as BET protein inhibitors. BET protein family contains 4 bromodomain containing proteins (BRD2, BRD3, BRD4, BRDT). These proteins were previously shown to be critical for AR-dependent gene expression. We detected interaction between BRD9 and BRD2/4. We demonstrated that BRD4 and BRD9 had shared binding sites on genome, a fraction of which are co-bound by AR. At particular target sites we showed that BRD9 localization is dependent on BET proteins, but not the other way around. Taking together, we provided some evidences that GBAF targeting through BRD9 can be a novel therapeutic approach for prostate cancer. Growing body of reports suggested that current therapy options targeting the androgen receptor is failing due to acquired resistance. Therefore, targeting the AR pathways via its coregulators such as BET proteins or SWI/SNF complexes can serve as potent alternative approaches. Further research is needed to elucidate the roles of GBAF and BET proteins in androgen receptor independent prostate cancer cells, which are still responsive to GBAF or BET manipulations although to a lesser extent.

CHAPTER 1.   INTRODUCTION

1.1   Basic concepts of gene expression regulation

One of the common features of all living things is responsiveness to environmental cues. For a simple bacterium, survival, fitness and reproduction rate are determined by the environmental parameters such as temperature, pH, electrolyte concentration, population density, nutrient availability or the presence of toxic substances. Changes in these parameters requires bacterium to adapt in order to survive or increase its fitness as much as possible.
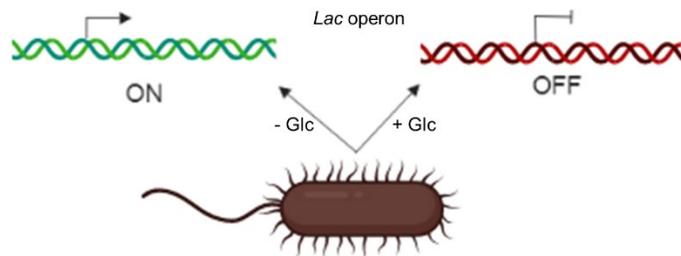
Figure 1-1 Any living creatures respond to intrinsic and extrinsic stimuli by altering the gene expression. A well-studied process is exemplified: When glucose (Glc) is limited in the environment, bacteria rewire the metabolism to be able to utilize alternative carbon sources such as lactose. This rewiring requires modifications in the gene expression profile, one of which is the activation of *lac* operon, which drives the expression of enzymes and transport proteins required for lactose internalization and catabolism. When the "favorite" nutrient glucose is available, the bacteria do not need to utilize metabolically-demanding lactose, thus the *lac* operon products. Therefore, the operon is kept inactive.

A detailed look at this "adaptation" process let us appreciate that the process is —at least partly— coded in the bacterium's genome —the total content of the DNA it possesses. When glucose, as the major carbon source, is scarce, bacteria seek to utilize alternative carbon sources such as lactose. This "sensing" process of the low glucose levels, availability of lactose and activating the required genes for lactose utilization form the example for one of the best understood means of prokaryotic gene regulation. A bacterium does not always express lactose-metabolism-related genes; rather, it "induces" when needed *i.e.* when glucose is limited (Figure 1-1).

Similar on/off circuitry or a "gradient" of expression is also seen in multicellular organisms, which is a little more complex than the regulation of gene expression in prokaryotes. From zygote to adult neurons, each cell type in humans contains the exact same DNA content, with certain

exceptions such as mature lymphocytes. However, functionally and phenotypically, a liver cell (hepatocyte) and neuron are different. When one tries to match the cellular activities in different cell types with their transcriptional profiles – or the total content of RNA products– we can appreciate the "specialization" at cell or tissue level. A hepatocyte does not conduct electrochemical signals in contrast to neurons; it does not "express" the proteins required for this type of signal transduction although it physically has the "genes" that code for them. Different tissues or organs and even certain cells within the same tissue differ in their major characteristics and this is very well reflected on their gene expression profiles. The specialization is connected to the set of the proteins or RNAs that a type of cell expresses, which indeed serves as the identity of the tissue or the cell types. Failure to establish or maintain this identity generally ends up with embryonic lethality, developmental disorders or malignancies (Figure 1-2).
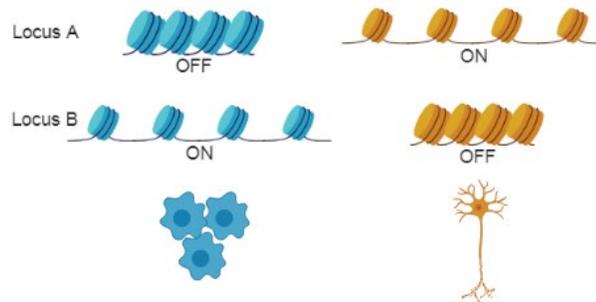
Figure 1-2 Tissue specification and lineage differentiation are initiated and maintained by non-genetic variations on the genome, also known as epigenetic changes. Epigenetic changes lead to alternative gene expression profiles within the same genome content, which does not involve changes in DNA sequence. The epigenetic regulations are key to formation of multicellularity, tissue specification and maintenance. Loosely or densely packed "beads-on-a-string" representation in the cartoon represents the major form of gene regulation occurring on eukaryotic genes, which will be discussed in detail in later sections.

As exemplified from both prokaryotes and eukaryotes gene expression regulation is a critical aspect of the life – from adapting to changing environmental conditions to evolution of multicellularity and lineage commitment. As a definition, regulation of gene expression comprises any mechanism that can alter the level or form of a gene product, that is, RNA or protein. These can be categorized as transcriptional, post-transcriptional/co-transcriptional (RNA splicing, RNA transport, RNA interference), translational and post-translational controls. This work will mostly focus on investigating the factors involved in transcriptional aspect of the gene regulation.

1.2 "Control tower" of the cell: Chromatin

In higher order eukaryotic systems DNA is packed into a protein-dense structure, called chromatin. It was initially thought that packing is critical to fit DNA, a huge macromolecule reaching 2 meters long per human cell, into the limited space of nucleus. Also, since DNA is vulnerable to damages, this packing is thought to have protective roles. With the advances in molecular biology and molecular genetics, the perception of "passive, protective" chromatin is challenged, the regulatory roles of the chromatin started to become prominent.
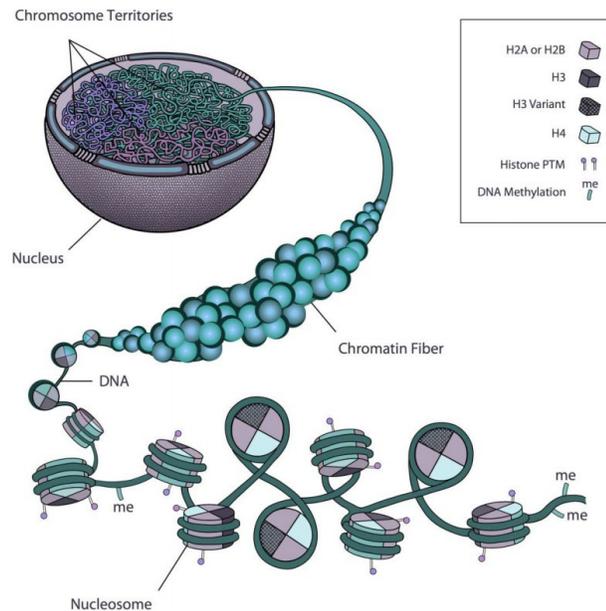
Figure 1-3 Hierarchy of chromatin assembly [1]. DNA is wrapped around octamer of small alkaline proteins called histones. DNA and the histone octamer form the basic repeating unit of chromatin, known as nucleosome, represented as "*beads* on a string" (Figure 1-2).

The basic unit of the chromatin structure is called a nucleosome. Nucleosomes consist of an octamer of 4 types of small basic proteins called histones (2 copies of each H2A, H2B, H3 and H4); and the DNA wrapped around this octamer. Another histone protein, H1 is not a part of the nucleosome; however, it plays key role in connecting nucleosomes and formation of 3D chromatin structure. Further packing into denser structures, nucleosomes form chromatin fibrils and eventually a 3D structure within the nucleus (Figure 1-3).

1.3　Gene expression regulation at the level of chromatin

Chromatin is a dynamic structure. The dynamicity qualifies the chromatin for controlling the cellular behaviors. As discussed earlier, from simple bacteria to higher-order eukaryotes, living creatures need to respond to extrinsic and intrinsic cues by turning on and off expression of certain genes. Besides, functional specialization in multicellular organisms such as tissues and organs require assigning multiple "meanings" to the same genetic code, which is unique to the individual organism. All these "alternative readings" of the code to form distinct functionalities are controlled at the level of gene expression; the chromatin has the pivotal role in this regulation.

In basic terms, the density of "packing" determines whether a gene segment will be read or kept silenced (Figure 1-2). When a chromatin region is densely packed, it is generally not accessible to the transcription machinery; therefore, it will have no or little transcription or "expression". These densely packed regions on chromatin are called heterochromatin regions. Expression from relatively open chromatin region is generally high since the DNA is accessible to transcription machinery. These loosely packed regions are called euchromatin. In normal cells, certain chromatin regions do not change their packing density. For instance, constitutive heterochromatin regions comprise centromeric or telomeric sequences or mobile elements, which are kept silent in *most* of the normal cell types [2]. Housekeeping genes, the genes that are essential for basic cellular functions, tend to be always in the form of euchromatin. However, most of the time, various cell types differ in their heterochromatin and euchromatin regions. During embryonic development or formation of highly specialized cell types from tissue-resident stem cells, the local chromatin state can switch between euchromatin and heterochromatin, leading to induction or repression of gene expression. Aberrant switches between on and off states are generally associated with diseases such as cancer.
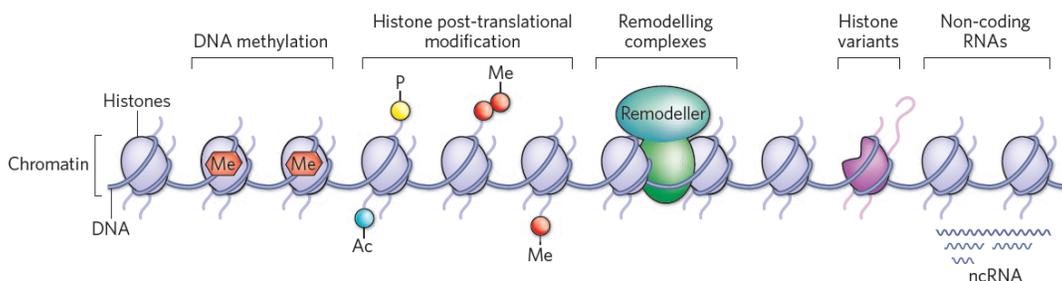
Figure 1-4 Factors regulating the chromatin structure, thereby enabling dynamic regulation of gene expression. Figure adapted from [3]

Chromatin structure can be altered by various processes [3]. These include DNA methylation, histone post-translational modifications, chromatin remodeling, histone variants and non-coding RNAs (Figure 1-4). These processes are interrelated; the crosstalk among various factors are critical for coordination of dynamic regulation of the chromatin structure, i.e. establishment, maintenance or disruption of a particular chromatin state.

DNA modifications and histone modifications are the most extensively studied mechanisms of chromatin regulation. DNA modifications includes covalent modifications on nucleobases, most characterized ones being the modification of cytosines. DNA methyltransferases (DNMT) are the enzymes that catalyzes the formation of 5-methylcytosines are generally associated with repressive transcription. However, the functional output of DNA methylation differs based on the genomic feature comprising the methylated cytosines, i.e. promoter, transcription start site, gene body, enhancer. For example, CpG-rich sites (also known as CpG islands) are generally under-methylated; still, methylation at CpG islands associated with transcriptional start sites linked to stable, long-term silencing as in the case of X-inactivation and imprinting [4]. However, cytosine methylation in gene bodies are not associated with gene repression. Besides, the same position on cytosines is subjected to alternative modifications, which are thought to be the "intermediates" of the demethylation reaction chain initiated by ten-eleven translocation (TET) dioxygenases [4,5]. The "demethylation" enzymes convert 5-meC to 5-hydroxymethylcytosine (5-hmC), a relatively stable modification on DNA. The subsequent oxidation products are 5-formyl-methylcytosine (5-fC) and 5-carboxy-methylcytosine (5-caC), which are less stable and eventually removed by base excision repair to yield unmethylated DNA.

Similar to DNA, nucleosomal histones are subjected to extensive post-translational modifications (PTMs) at both unstructured N-terminal tails or at the cores. These modifications have roles in overall chromatin structure modulation and regulation of gene expression, DNA repair or replication. Histones can undergo methylation, phosphorylation, acetylation, ADP-ribosylation, small protein modifications (ubiquitination, SUMOylation), glycosylation, various understudied acylations (butyrylation, crotonylation, succinylation) and serotonylation [5–8]. Apart from the biological function of the modification *per se,* histone modifications have different outputs based on the modified residue, the valency of modification (e.g., mono-, di-, tri-methylation) and the feature associated with the modification. The extensive repertoire of

modifications and the vast number of modifiable residues further expands the functional outputs from histone modifications by means of "combinatorial readout" [9].

1.4   ATP-dependent chromatin remodelers

Another group of factors that are involved in dynamic regulation of chromatin structure is the ATP-dependent chromatin remodelers. In principle, the chromatin remodeling is collectively categorized into three major activities: Assembling the nucleosomes on DNA; altering the physical spacing between the nucleosomes by moving them along DNA or evicting them; exchanging the nucleosomal histones with their variants [5,10,11]. All activities require hydrolysis of ATP by the ATPase –the chromatin remodeler itself, which is utilized to sequentially disrupt (while repositioning) or form (while assembling) fourteen histone-DNA contacts that enables *freeing the DNA as a "wave" emerging from the surface of DNA (while repositioning)* or *bending the DNA over nucleosome surface (while assembling)*.
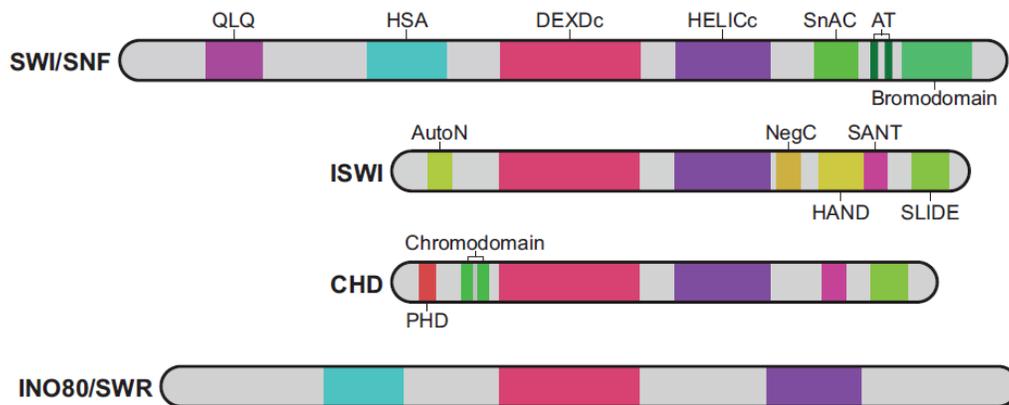
Figure 1-5 Domain structures of four classes of ATP-dependent chromatin remodelers: Switch/Sucrose non-fermentable (SWI/SNF); imitation switch (ISWI); chromatin helicase DNA-binding (CHD); INO80. ATPase domain, which is composed of DEXDc and HELICc lobes, is common to all chromatin remodelers. Figure adapted from [11]

There are four classes of ATP-dependent chromatin remodelers that differ in their domain structures and activities (Figure 1-5). The ATPase domain that contains DEXDc and HELICc lobes are conserved among all classes [10,11]. The unique activities of each remodeler are mostly regulated by their specific domains and the associated subunits. Briefly, nucleosome assembly is mostly fulfilled by ISWI and CHD remodelers that mediate random histone deposition and maturation of

nucleosomes [10], which can also facilitate gene repression. Accessibility of naked DNA is mostly adjusted by SWI/SNF type of remodelers: They can slide the nucleosomes, eject mature octamers or dimers from the chromatin, which enables nucleosome-free DNA accessible to other factors acting on chromatin such as sequence-specific transcription factors. INO80 type of remodelers are generally associated with incorporation and eviction of histone variants such as H2AZ whose dynamic exchange is critical during both transcriptional activation and repression as well as DNA damage repair [10,12] (Figure 1-6).
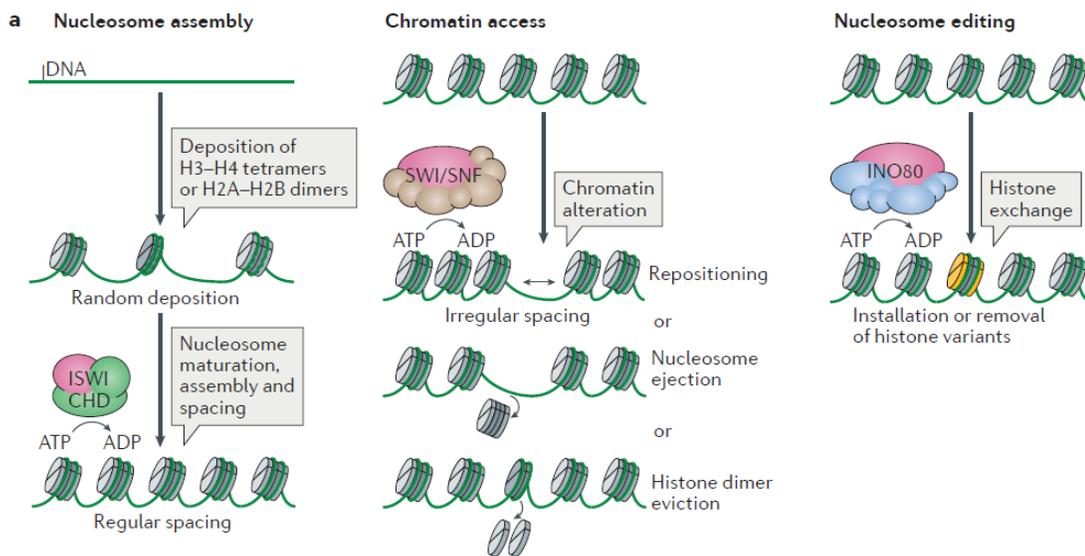
Figure 1-6 Summary of distinct mode of actions for ATP-dependent chromatin remodelers. Figure adapted from [10]

1.5 SWI/SNF complexes in development and disease

SWI/SNF complexes are one of the four major classes of mammalian ATP-dependent chromatin remodelers. The complexes have been discovered in yeast when the factors required for mating type switching upon mating factor signaling were screened [13]. Similar to other chromatin remodelers, SWI/SNF complexes are also multisubunit assemblies that contain core ATPase and several accessory proteins that are involved in regulation of remodeling activity, genomic targeting via interaction with transcription factors, DNA or histone marks and integrity of the complex [14,15]. In the course of evolution of multicellularity, new subunits were added into SWI/SNF complexes, while most of the yeast subunits were still conserved. In vertebrates, however, both the number of subunits and the number of paralogs per subunits were increased, leading to formation of highly

diverse subcomplexes by means of combinatorial assembly. This diversification is critical for ensuring tissue- and developmental stage- specific functionalities [5]. It is estimated that 29 different genes code for 15 SWI/SNF subunits within mammalian SWI/SNF complexes, which is subject to changes depending on tissue and developmental stage [13].

Mammalian SWI/SNF complexes (or BAF complexes) can assume different architectures. Until recently, two major types of SWI/SNF complexes have been described in mammals: BAF (or cBAF, for canonical BAF) and PBAF (polybromo-BAF). The major distinction between these complexes is the subunits that are unique to either complexes (Figure 2-1). For instance, although majority of the subunits are common to both, PBAF exclusively incorporates ARID2, BRD7, Polybromo-1 and BAF45A while only BAF incorporates BAF45B/C/D, SS18 or BCL7A/B/C. Apart from the architectural classification, SWI/SNF complexes can be categorized as tissue- or developmental stage-specific forms, which is mainly regulated at the level of expression of select paralogs. For example, BAF53B is the defining subunit of neuronal BAF (nBAF), which is the prevalent form of BAF complexes in post-mitotic neurons; while BAF53A is more universal [16]. Similarly, BAF60C is mostly incorporated in cardiac-specific BAF complexes [17,18].

Among the four classes of ATP-dependent chromatin remodelers, SWI/SNF complexes have a unique biological feature: Studies showed that more than 20% of all cancers have mutations in SWI/SNF subunit genes [5,13,19], being one of the most commonly mutated factors across different cancer types. The comprehensive analysis of the genome- and exome-sequencing of various cancer types indicated that mutations in specific SWI/SNF subunits were observed only in specific types of cancers [19]. Likewise, paralogous subunits have different mutation frequencies. These observations implied several critical points: First of all, despite functioning as a part of multimeric complex, not all subunits have the same degree of significance for a given tissue type. For instance, BAF47 is one of the core subunit of both BAF and PBAF assemblies; it has been shown to be involved in engagement of the complex with the nucleosomes and regulation of remodeling activity [20,21]. Its biallelic loss is observed approximately in 100% of malignant rhabdoid tumors (MRT) *as one of the few overall mutational events associated with malignant rhabdoid tumors* while BAF47 is rarely mutated in other cancer types, suggesting a unique sensitivity of MRT formation to BAF47 loss. Similarly, BCL7 and BCL11 paralogs are mostly mutated in hematological cancers but not in solid tumors [19]. Taking together, it is possible that both at the level of complex and individual subunits, different tissues or developmental stages have different

dependencies to SWI/SNF complexes. These context-specific roles of the complexes can only be elucidated by studying individual subunits and individual complex configurations and by integration of tissue-specific factors.

In certain instances, subunit inactivation can lead to therapeutic opportunities by means of synthetic lethal interactions among the complex members. This becomes more evident when one of the paralogs of a subunit family is inactivated. For instance, mammalian SWI/SNF contains one of the two ATPase subunits, BRG1 or BRM, which are usually co-expressed in different tissues. BRG1 is the one of the most highly mutated SWI/SNF member across the cancer types, whereas BRM mutations are rare [19]. It was shown that BRG1-deficient cells are sensitive to BRM-depletion in different cancers, suggesting that loss of BRG1 functionality -at least partially- is substituted by BRM [22]. Likewise, BRM-deficient esophageal squamous cell carcinoma cells are more dependent on BRG1 function [23]. Similar synthetic lethal interactions have been reported for other mutually exclusive BAF subunits such as ARID1A-ARID1B: ARID1A mutations in ovarian cancer and colorectal cancers render the cells dependent on ARID1B, a subunit that has lower expression than ARID1A in most of the tissues [19,24,25]. Mechanistically, it was demonstrated that the regulation of the chromatin accessibility by ARID1A/1B is the key reason for the synthetic lethal interaction [26]. Although ARID1B has little effect in chromatin accessibility in ARID1A wild type cells, in the absence of ARID1A, ARID1B partially substitute ARID1A-specific sites, implying a potential reason for synthetic lethality.

1.6   The novel, recently-identified SWI/SNF subcomplex: GBAF

GLTSCR1/1L-BAF (GBAF) subcomplex (also reported as non-canonical BAF or ncBAF [27]) has recently been defined compositionally and functionally in several studies including ours [27–31]. These studies took similar approaches in different cell line systems including various human cancer cell lines, normal immortalized cell lines and mouse embryonic stem cells to define the composition of the complex. Consistently, previously known BAF subunits such as BRG1, BAF155, BAF60, BCL7, SS18, ACTIN, BAF53A were incorporated into the complex. Interestingly, the subcomplex lacked some other core subunits such as ARID1A/B, BAF170, BAF47, BAF57, BAF45. The complex, on the other hand, uniquely incorporate an uncharacterized protein named glioma tumor suppressor candidate region 1 (GLTSCR1), its putative paralog GLTSCR1L and bromodomain-containing protein 9 (BRD9).

Earlier high-throughput proteomics studies identified GLTSCR1 as a BAF member [32]; however, this studies did not address the context in which GLTSCR1 is a BAF member. Using targeted co-IPs, we and others defined that GLTSCR1 is not a part of known BAF configurations, canonical BAF (BAF or cBAF) and Polybromo-BAF (PBAF) since it did not coimmunoprecipitate BAF47, BAF170, ARID or Polybromo-1 as both common and dedicated subunits of BAF and PBAF. GLTSCR1L, however, even more mysterious protein. It appears that the paralogs are mutually exclusive and overexpression of one paralog led to reduction in the protein level of the other, possibly through outcompeting it for complex incorporation [28]. It is also reported that during differentiation of mouse embryonic stem cells, GLTSCR1 goes up while GLTSCR1L goes down, suggesting differential functionalities of the complexes [29]. **In fact, *GLTSCR1L* knockout mice is embryonic lethal, while *GLTSCR1* knockout phenotype has not been reported.** Both proteins share a C-terminal conserved region that is named as "GLTSCR domain". A deletion experiment suggested that construct that lacks the C-terminal conserved domain cannot interact with GBAF subunits, suggesting that the conserved domain is critical for complex incorporation [27]. Further studies are needed to define the functions of extra-C-terminal regions. Indeed, in a tiled CRISPR screening for *GLTSCR1* gene to identify the *most critical regions* of the gene for the overall fitness, the major cytotoxicity is conferred by the guide RNAs targeting the C-terminal domain [27]. The assay implied that the errors in this region is not well- tolerated.

BRD9 is a relatively well-characterized dedicated subunit of GBAF. As a unique and non-redundant subunit of GBAF, it can serve as a probe to study GBAF functions. Before recognition of BRD9 as GBAF subunit, it was shown to be a critical driver of *MYC* expression in acute myeloid leukemia cells, which are particularly sensitive to BRD9 depletion [33]. Another study demonstrated its antagonistic relationship with highly similar bromodomain containing protein BRD7 that incorporates to PBAF subunit [34]. It was shown that vitamin D-mediated transcription is held repressed when BRD9 engages an acetyl-lysine residue on vitamin D receptor (VDR). Upon ligand induction of VDR the same residue is engaged by BRD7, which enables open chromatin and VDR-mediated transcription. Functionally, the study demonstrated that active VDR signaling is important for pancreatic β cell survival and repression of hyperglycemia under type 2 diabetes and offering BRD9 targeting as a viable pharmacologic approach.

Recent studies emphasized the dependence to GBAF in SWI/SNF-defective cells. BAF47 is one of the core SWI/SNF subunit that is critical for remodeling activity [21]. *SMARCB1*, the gene

encoding BAF47, is biallelically inactivated in rhabdoid tumors [19,20,35]. Two recent studies showed that *SMARCB1*-null cells are exclusively sensitive to *BRD9* depletion [27,31]. In another case, a rare soft-tissue cancer synovial sarcoma contains fusions of SS18-SSX that form aberrant BAF complexes. The fusion protein assembles into BAF complex in place of wild type SS18 but blocks BAF47 incorporation [36]. Similar to malignant rhabdoid tumors, synovial sarcoma cells with the fusion proteins are sensitive to *BRD9* depletion [27,30]. Interestingly, depletion of both SS18-SSX as disease driver and BRD9 attenuated the growth; however, transcriptional effects of both manipulations are only concordant at cell-cycle genes, suggesting that they act through different targets, which suggested that GBAF and BAF functions are not completely redundant. Overall, despite foreseen functional redundancies, different subcomplexes may serve as unique dependencies in specific conditions. This necessitates studying the complexes with both SWI/SNF-intact and SWI/SNF-defective settings.

1.7 Summary

Gene regulation is a critical component of cellular homeostasis in both eukaryotes and prokaryotes, although the complexities of the regulatory machines differ significantly between the two domains, which makes perfect sense comparing their genome complexities in terms of both number of genes and amount of non-coding regulatory sequences. In the complicated nature of multicellularity and maintenance of the harmony within and among all distinct types of tissues, a reasonable understanding of gene expression regulation at the level of chromatin structure is necessary. Studies have associated many types of disease such as developmental defects and malignancies with aberrations in the machineries acting on the chromatin. To this end, we focused on one of the major class of chromatin regulators, SWI/SNF complexes. We discovered an alternative type of SWI/SNF complex and tried to assign a functional role in development and disease state. We are hopeful that this and similar studies can lead to better understanding of human diseases and help discovery of novel therapeutics.

CHAPTER 2.   BIOCHEMICAL CHARACTERIZATION OF GLTSCR1/1L-BAF (GBAF) COMPLEX

2.1   Introduction

The mammalian SWI/SNF (or BAF) complex is an ATP-dependent chromatin remodeler composed of 10-14 subunits [37]. The mammalian SWI/SNF chromatin remodeling complex is implicated in a variety of processes including mitosis, DNA replication, DNA damage repair, genomic looping, gene splicing, in addition to its well-established roles in the transcriptional regulation of genes involved in cellular differentiation, cellular maintenance, and adaptation to stimuli [38]. Mutations in specific SWI/SNF complex members are common in cancer [39,40] and neural disorders [41], and the altered expression of specific subunits is associated with tumorigenesis [42], viral infection [43], viral latency [44], alcohol addiction [45,46], heart disease [47], and immune function [48]. The ability of this complex to direct such numerous and diverse functions is facilitated through the increase in SWI/SNF subunit number and diversity during vertebrate evolution [49], which led to an exponential increase in the potential combinations of subunits [50,51]. All SWI/SNF complexes contain the ATPase subunit BRG1 or BRM, along with the structural subunits BAF155/BAF170, which are required for full ATPase and nucleosome remodeling activity in vitro [52]. In addition, SWI/SNF complexes contain BAF60 (A, B or C), BAF47, BAF57, BAF53 (A or B), and actin. The larger and less abundant PBAF complex, uniquely contains ARID2, PBRM1, BAF45d and BRD7, while the more abundant BAF complex contains ARID1 (A or B), BAF45 (B, C, or D), SS18, BCL7 (A, B or C), and BCL11 (A or B) **(**Figure 2-1**).** The altered expression of SWI/SNF

paralogs during cellular differentiation results in subunit switching, which is an important determinant of cell identity and cell-type transcriptional programs [53].
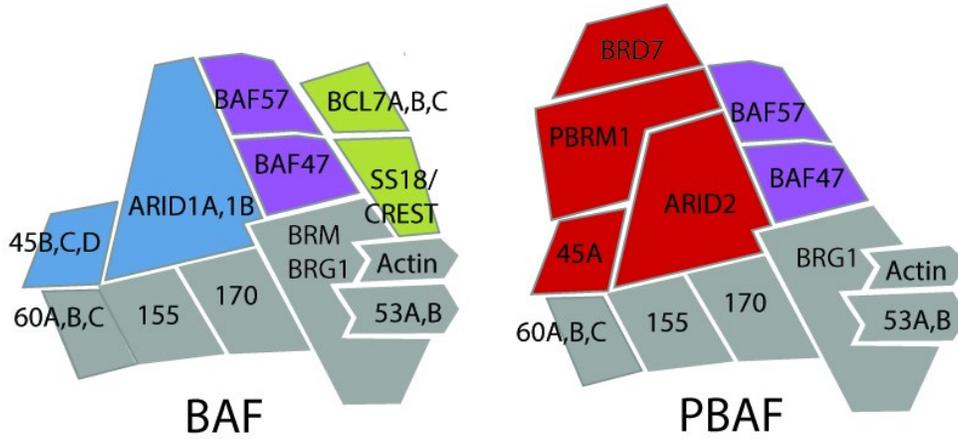
Figure 2-1 Current illustration of mammalian SWI/SNF complex composition

Additionally, paralogs are often expressed simultaneously, leading to distinct subcomplexes within the same cell with both unique and redundant function [32]. For example, ARID1A is high in embryonic stem cells while ARID1B is upregulated upon differentiation [54]. The different BAF complexes containing these two paralogs share many of their genomic targets; however, they also bind unique genomic targets and deletions are non-synonymous for gene regulation [55]. ARID1A is the most commonly mutated SWI/SNF subunit in cancer, due to transcriptional functions that are non-redundant with ARID1B [26,56]; however, cancers with deletions in ARID1A are dependent on ARID1B for viability [24] due to redundant, essential functions at enhancers [26]. Additionally, homologous complexes can display transcriptionally antagonistic roles, as has been observed for ARID1A and ARID2-containing complexes at specific gene targets [44,55,57]. Targeting specific SWI/SNF complexes has been proposed both for alleviating subunit-specific pathogenic function as well as to target essential redundant functions in cancers with mutations in the genes for specific subunits [58,59]. Both strategies are dependent on a better understanding of the different biochemical and transcriptional functions of homologous SWI/SNF complexes. We report here for the first time a novel, ubiquitously expressed SWI/SNF subcomplex defined by mutually exclusive paralogs GLTSCR1 (or BICRA for BRD4-Interacting Chromatin Remodeling Complex Associated) and GLTSCR1L (or BICRAL for BRD4-Interacting Chromatin Remodeling Complex Associated Like), which also contains BRD9 and a subset of shared canonical SWI/SNF subunits.

2.2   Results

Proteomic analysis of BRG1 immunoprecipitations from two renal clear cell carcinoma cell lines identified multiple unique peptides from the uncharacterized protein GLTSCR1 (Figure 2-2A).
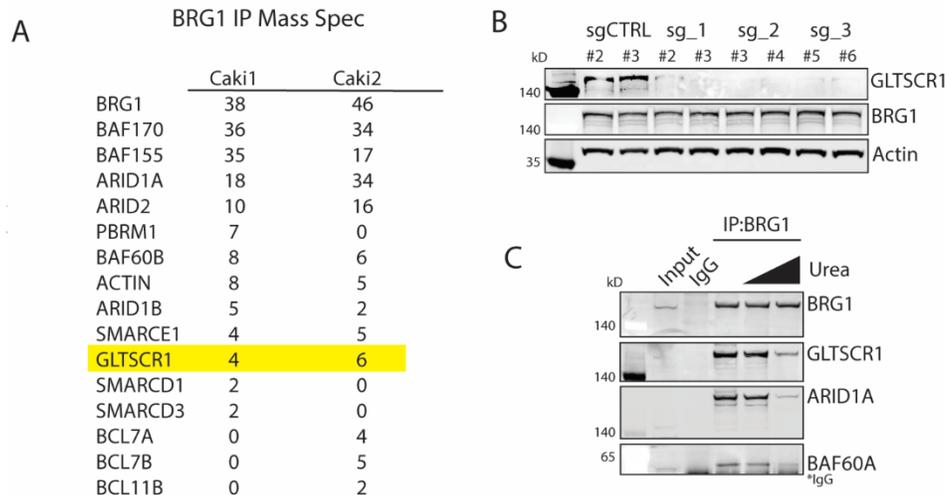
Figure 2-2 GLTSCR1 is a dedicated subunit of the SWI/SNF chromatin remodeling complex. A) Mass spectrometry analysis of BRG1 IP from two human renal cancer cell lines identifies peptides from GLTSCR1. B) GLTSCR1 specific antibody identified using Gltscr1 knockout mESCs derived using three different sgRNA constructs. C) Immunoprecipitation with antibodies against BRG1 confirms GLTSCR1 association. Urea denaturation with 0.5 M and 2.5 M urea prior to BRG1 immunoprecipitation indicates the strong association of GLTSCR1 to BRG1, comparable to the strength of association of BRG1 to core BAF subunits ARID1A and BAF60A.

GLTSCR1 has been identified in previous proteomic analyses of the SWI/SNF chromatin remodeling complex [32,60–62] but has never been validated or characterized as a BAF complex subunit. After screening multiple commercially available antibodies against GLTSCR1, we identified an antibody that stained a band in the predicted region of 180 kDa using immunoblot analysis. Further, this band disappeared after CRISPR-mediated Gltscr1 knockout in mouse embryonic stem cell lines (Figure 2-2B). Using this validated antibody, we confirmed the mass spectrometry data using immunoblot analysis, detecting robust enrichment of GLTSCR1 in BRG1 immunoprecipitations (Figure 2-2C). To define if GLTSCR1 is a true subunit of BAF and not an associating factor, we performed urea denaturation followed by BRG1 immunoprecipitation and

found that GLTSCR1 stably associates with BRG1 at urea concentrations up to 2.5 M, consistent with known BAF subunits ARID1A and BAF60A (Figure 2-2C).

To determine which SWI/SNF subcomplex contains GLTSCR1, we performed glycerol gradient analysis to separate the two closely related SWI/SNF complexes, BAF and PBAF, based on density. Surprisingly, GLTSCR1 staining was detected in earlier gradient fractions 11-13, which did not overlap with ARID1A (a subunit exclusive to the BAF complex) in fractions 14-16 or PBRM1 (a subunit exclusive to the PBAF complex) in fractions 17-19 or with BAF47 (a subunit shared by both BAF and PBAF) (Figure 2-3A-left). To ensure that this was not an aberrant partial complex due to specific cancerous alterations or cell culture artifacts, we performed similar analysis in a second cell line (PC3) and observed the same pattern for GLTSCR1 (Figure 2-3A-right). To define if any additional SWI/SNF subunits besides BRG1 associate with this subcomplex, we performed a series of immunoprecipitations to various known subunits of the BAF or PBAF complex (Figure 2-3B). From this panel, only antibodies against BRG1 and BRM were able to precipitate GLTSCR1, and, as expected from the glycerol gradient analysis, GLTSCR1 did not associate with BAF-specific subunits ARID1A/ARID1B or PBAF specific subunit PBRM1. Surprisingly, though, we did not observe GLTSCR1 association with BAF45D or BAF57, subunits thought to be canonical subunits, although we did observe association with core subunits BAF155 and BAF53a (Figure 2-3C). Using GLTSCR1 immunoprecipitations in HEK293T cells, we further identified that BAF60A, SS18 and BRD9 are GBAF subunits while BAF170 and BAF47 are not (Figure 2-3D). Further, glycerol gradient analysis and co-immunoprecipitation experiments identified BRD9 as a subunit exclusive to GBAF, SS18 as a subunit shared by BAF and GBAF, and BAF170 and BAF47 as subunits exclusive to BAF (Figure 2-3E, Figure 2-3F). An illustration of the proposed composition of these complexes based on the immunoprecipitation experiments is depicted in Figure 2-3G. To validate GBAF as a potential chromatin remodeling complex, we next performed ATPase assays on immunoprecipitations of GLTSCR1 and BRG1 from HEK293T cells. We used GLTSCR1 immunoprecipitations containing comparable amounts of BRG1 (~90%) (Figure 2-3H) and found that GBAF complexes display robust DNA-stimulated ATPase activity (Fig 2H). In fact, the ATPase activity of GLTSCR1 immunoprecipitations was higher than BRG1 immunoprecipitations, although this is complicated by possible contributions from BRM, which is lowly expressed in HEK293T cells [63]. We next used sequential salt extraction assays and determined that even in the absence of association with DNA-binding subunits BAF57 and

ARID1/2, GLTSCR1 elutes from bulk chromatin with similar salt concentrations as BAF-specific subunit ARID1A, while PBAF-specific subunit PBRM1 requires slightly higher salt concentrations to elute off bulk chromatin [64] (Figure 2-3I).

Figure 2-3 GLTSCR1 is in a novel SWI/SNF subcomplex GBAF. A) Glycerol gradients from renal cancer cell line Caki1 and prostate cancer cell line PC3 indicate that GLTSCR1 does not cosediment with BAF subunit ARID1A or PBAF subunits PBRM1 (for Caki1) or BRD7 (for PC3). B) IP experiments of BAF subunits from PC3 lysates identify GLTSCR1 association with BRG1 and BRM. C) BAF subunit and GLTSCR1 IP experiments from HEK293T lysates identify GLTSCR1 association with BAF155 and BAF53a but not BAF47, BAF57 or BAF45D. D) BAF subunit and GLTSCR1 IP experiments from HEK293T lysates identify GLTSCR1 association with SS18 and BRD9 but not BCL11A. E) Glycerol gradient analysis and F) BAF subunit IP experiments from HEK293T lysates identify GLTSCR1 association with BRG1 and SS18 but not BAF170 and BAF47, and validate BRD9 as a subunit exclusive to GBAF. G) Schematic representation of GBAF, BAF and PBAF composition. Yellow subunits are unique to GBAF, blue subunits are unique to BAF, red subunits are unique to PBAF, green subunits are shared by GBAF and BAF, purple subunits are shared by BAF and PBAF, and grey subunits are shared by all three complexes. Subcomplex GBAF consists of BAF60A, BRG1, BAF155, BRD9, BAF53A and SS18. H) GBAF possesses ATPase activity. ATPase activity assay was performed with BRG1 and GLTSCR1 immunoprecipitations providing similar levels of BRG1. ATPase activities normalized to respective IgG isotype controls yielded comparable fold changes [3.03 ± 0.23, for BRG1 IP; 3.24 ± 0.87, for GLTSCR1 IP]. Error bars, mean ± s.d. $n = 3$. * $p<0.05$, *** $p< 0.001$; I) Sequential salt extraction analysis and immunoblot quantitation indicates that GLTSCR1 interacts with bulk chromatin at a similar strength as ARID1A (representative of BAF), and PBRM1 (representative of PBAF).

31
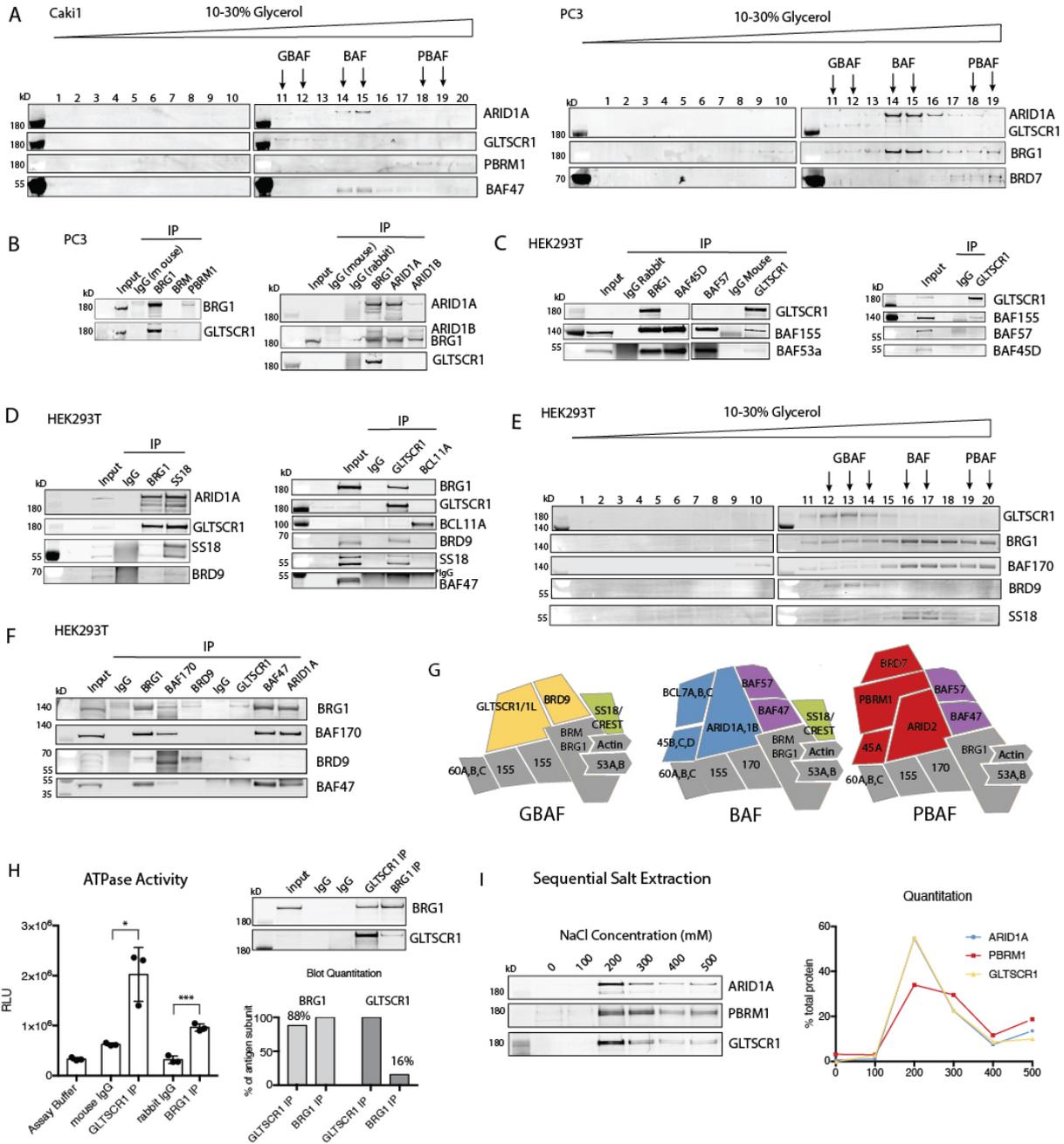
As we established GLTSCR1 as the unique subunit of GBAF, we set to define whether GLTSCR1 is required for GBAF formation. Using GLTSCR1 knockout ES cells (Figure 2-2B) we performed glycerol gradients with and without GLTSCR1 (Figure 2-4A). We observed a decrease in BRG1 and BAF60A staining in GBAF fractions 11-13, but not complete loss of staining. We hypothesized that this was due to the presence of the predicted GLTSCR1 paralog, GLTSCR1L (now referred to as BICRAL for clarity), which has also been detected in BAF subunit IP mass spectrometry studies as KIAA0240 [61,62]. GLTSCR1 and BICRAL share 32% sequence homology (21% identity) and both contain a well-conserved "GLTSCR1" domain, which is also conserved between GLTSCR1 orthologs predicted in all multicellular organisms (Figure 2-4B). We screened commercially available antibodies for BICRAL and identified one with weak staining at the predicted size of 140 kDa, along with many non-specific bands. To confirm that the band is the correct protein, we developed a cell line with doxycycline-inducible FLAG-tagged BICRAL. Overexpression of BICRAL-FLAG in HEK293T cells resulted in a robust FLAG band at 140 kDa and an increase in staining with the endogenous antibody at the same molecular weight (Figure 2-4C). To confirm that BICRAL is mutually exclusive with GLTSCR1 in the GBAF complex, we performed co-immunoprecipitations of GLTSCR1 and FLAG in our BICRAL-FLAG overexpression system and using endogenous proteins in HEK293T cells, and found that GLTSCR1 and BICRAL do not associate with each other (Figure 2-4D, Figure 2-4E). In addition, both GLTSCR1 and BICRAL-FLAG enrich BRG1, BAF60A, and BRD9 but not ARID1A, indicating incorporation into comparable SWI/SNF subcomplexes. Intriguingly, we also found that overexpression of BICRAL decreases GLTSCR1 expression, possibly indicating its ability to compete with, and replace GLTSCR1 in GBAF complexes.
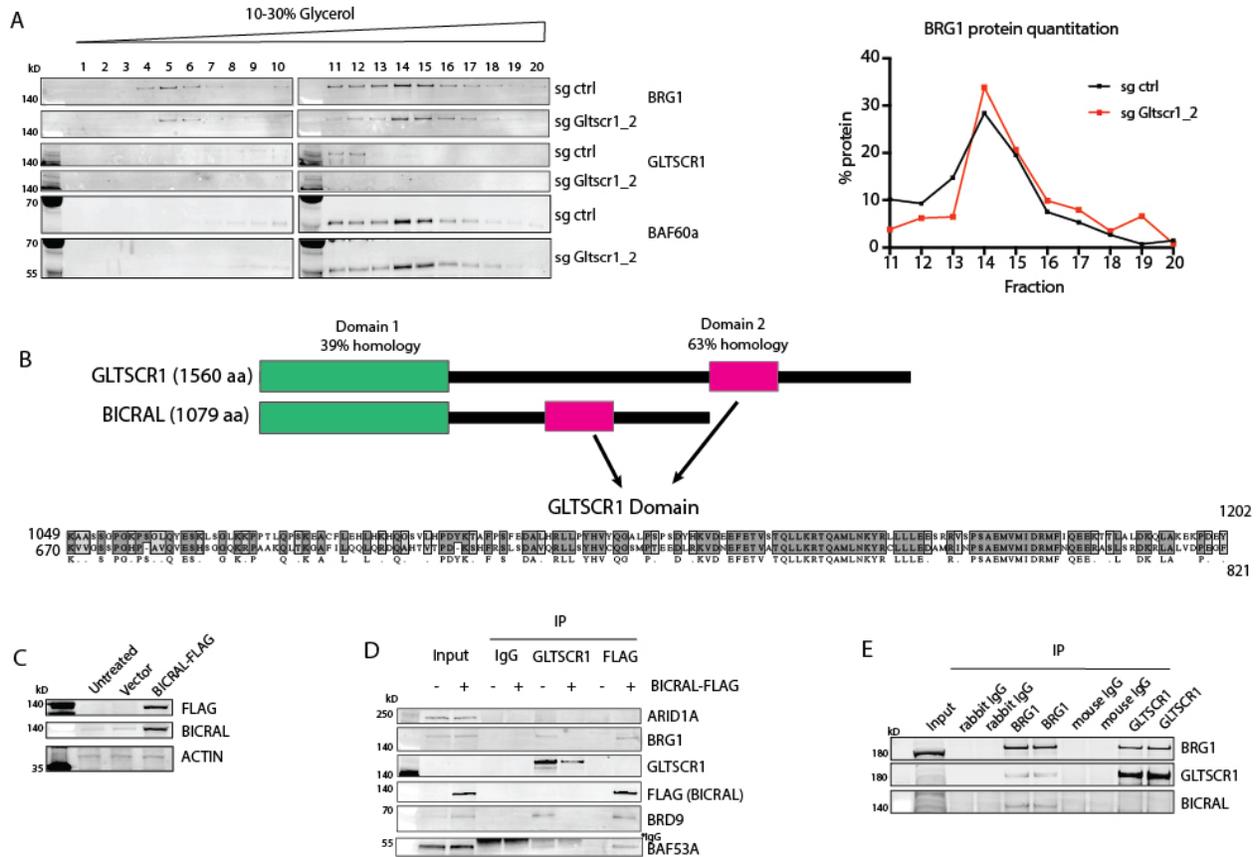
Figure 2-4 GBAF contains GLTSCR1 or paralog GLTSCR1L (BICRAL). A) Glycerol gradient analysis in mESCs showing that GBAF-associated BRG1 and BAF60A were preserved in fractions 11-13 in the absence of GLTSCR1, suggesting that GBAF was not completely disrupted by Gltscr1 knockout. B) Pairwise alignment of amino acid sequences of GLTSCR1 and its paralog GLTSCR1L (BICRAL) show homology in the N-terminal region and strong homology at region identified at a conserved GLTSCR1 domain C) Verification of inducible expression of BICRAL-FLAG in HEK293T cells with both FLAG and endogenous BICRAL antibodies. D) Immunoprecipitation analysis showed that similar to GLTSCR1, exogenous BICRAL interacts with BRD9, BAF60A and BRG1. In addition, endogenous GLTSCR1 and BICRAL do not immunoprecipitate one another. BICRAL overexpression results in reduced GLTSCR1 protein levels. E) Endogenous BICRAL does not associate with GLTSCR1, further validating GLTSCR1 and BICRAL are mutually exclusive in GBAF context. BICRAL is detected in total BRG1 IP but not in GLTSCR1 IP, although both contain comparable levels of BRG1.

To further investigate the role of BICRAL in GBAF formation we performed glycerol gradient analysis of BICRAL-FLAG overexpression in HEK293T cells. We found that BICRAL overexpression results in BICRAL incorporation into GBAF, as indicated by its expression in fractions 11-13, similar to the profile of GLTSCR1 staining (Figure 2-5A). In addition, we

confirmed that it is able to replace GLTSCR1 in GBAF, as indicated by an overall decrease in GLTSCR1 staining. Interestingly, we also saw an increase in BRG1 staining in fractions 11-13 upon BICRAL overexpression, indicating that BICRAL is able to alter overall SWI/SNF complex stoichiometry. To investigate this further, we performed BRG1 immunoprecipitations in the BICRAL overexpression cells. We found no changes in BRG1 expression or immunoprecipitation efficiency and confirmed the decrease in GLTSCR1 association with BRG1. In addition, we observed an increase in both the expression and the BRG1 association of BRD9 and a decrease in the BRG1 association with BAF subunits ARID1A, BAF47, and BAF57 (Figure 2-5B). To test whether this effect is due to a transcriptional or post-translational outcome of BICRAL expression, we performed RT-qPCR in BICRAL-FLAG overexpression line and found no alterations in endogenous transcript levels for any of the subunits in question (Figure 2-5C). This suggests that increased BRD9 protein levels and decreased GLTSCR1 levels are due to post-translational events, most likely degradation of free monomer. To confirm that GBAF is dependent on GLTSCR1 or BICRAL for formation, we also generated a double knockout HEK293T cell line. We observe that BRG1-associated BRD9 is undetectable in GLTSCR1 knockout and double knockout cells, indicating loss of GBAF formation (Figure 2-5D). It is worth noting that differences in knockout efficiencies and possibly the relative levels of GLTSCR1 and BICRAL made it difficult to distinguish additive or GLTSCR1-dominant effects of the paralogs on GBAF formation. Similar to decreased GLTSCR1 levels in BICRAL-overexpression lines, we consistently observed an increase in BICRAL levels upon GLTSCR1 knockout (Figure 2-5D), via a similar increase of protein stability through complex incorporation. This provides further evidence for a compensatory role of BICRAL for GBAF formation in the absence of GLTSCR1. These results indicate that GLTSCR1/BICRAL are mutually exclusive subunits of GBAF that can, in part, define SWI/SNF complex stoichiometry.
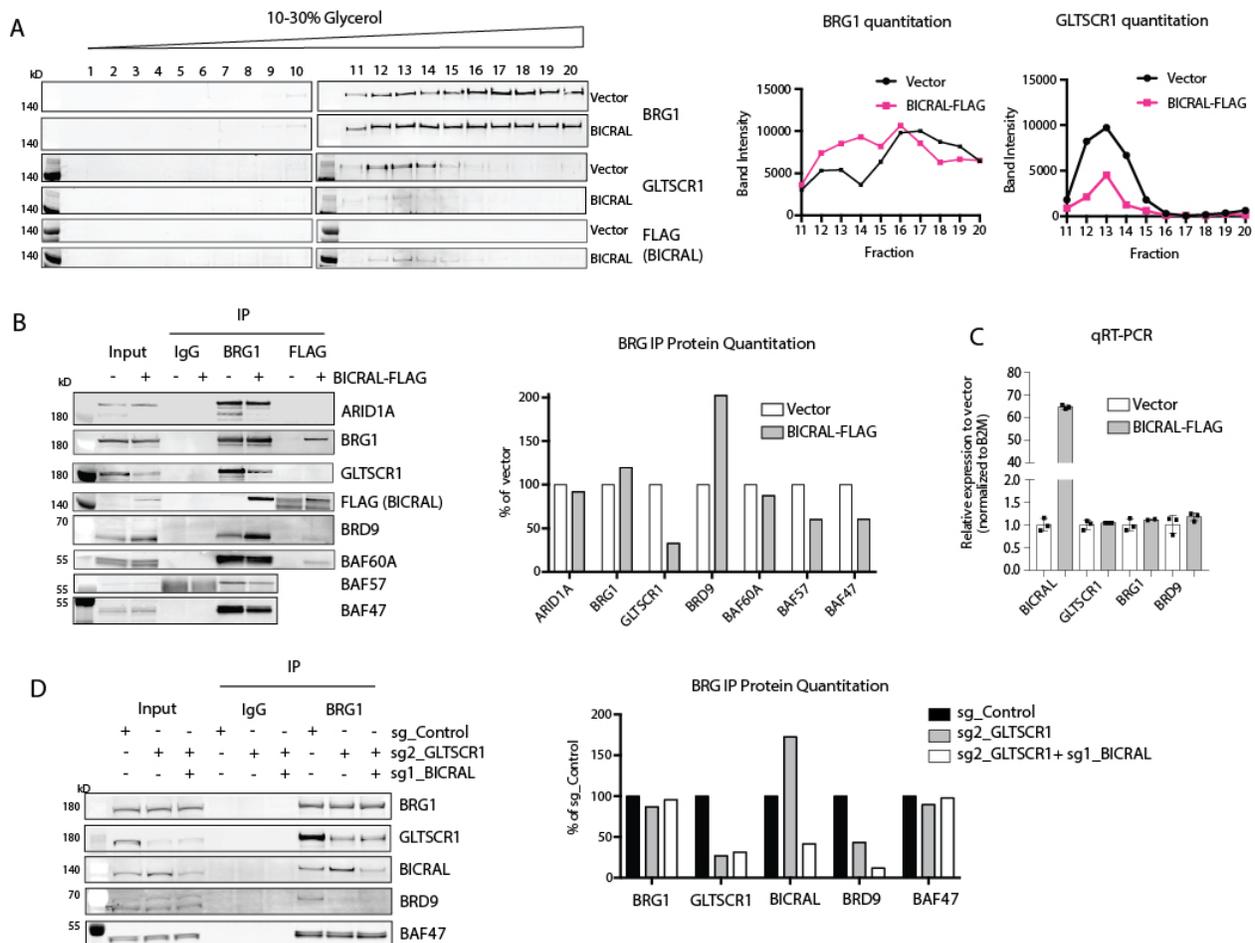
Figure 2-5 GLTSCR1 and BICRAL are mutually exclusive subunits of GBAF that can alter SWI/SNF complex stoichiometry. A) Glycerol gradient analysis in BICRAL-FLAG-overexpressing HEK293T cells indicates that BICRAL is incorporated into GBAF. Overexpression of BICRAL-FLAG increases the GBAF-associated BRG1 levels (fractions 11-14) suggesting formation of new GBAF upon BICRAL overexpression. Reduced GBAF-associated GLTSCR1 levels also validates decreased GLTSCR1 protein expression upon BICRAL overexpression. B) Immunoprecipitation analysis showing that BICRAL-FLAG overexpression reduced BRG1-associated GLTSCR1 levels and enhanced BRD9 protein levels and its association with BRG1. BICRAL-FLAG overexpression also reduced BRG1-associated BAF47 and BAF57, suggesting competition between GBAF and BAF for BRG1. C) RT-qPCR showing that expression of BRG1, GLTSCR1 or BRD9 did not alter upon BICRAL overexpression. Error bars, mean ± s.d. n = 3. D) CRISPR/Cas9-mediated knockout of GLTSCR1 with or without CRISPR/Cas9-mediated knockout of BICRAL reduced the BRG1-associated BRD9 levels, as an indicator of loss of GBAF.

GLTSCR1 has also been identified in a proteomics study of BRD4-associating factors [65], as reflected by the recent change in HUGO gene name from GLTSCR1 to BICRA for BRD4-

Interacting Chromatin Remodeling Complex Associated protein. The BRD4 extraterminal (ET) domain was found to associate with several proteins, including NSD3 (and NSD2), ATAD5, GLTSCR1, and CHD4 (and CHD7), in an ET domain-specific manner [66]. Using BRD4 immunoprecipitations we confirmed that BRD4 associates with GLTSCR1, BAF155, BRD9, BAF60A but not BAF-specific subunit BAF47 (Figure 2-6A). Since BRD4 protein association and coregulator activity are known to be regulated by phosphorylation [67], we treated BRD4 immunoprecipitation with alkaline phosphatase (or used phosphatase inhibitors in lysis, IP and washes) but did not find that the association between GLTSCR1 and BRD4 is dependent on phosphorylation. We confirmed previously published findings that androgen-sensitive prostate cancer cell line LNCaP is ten-fold more sensitive to BRD4 inhibition than androgen-insensitive prostate cancer cell line PC3 [68]; however, LNCaPs are not dependent on GLTSCR1 for viability (Figure 2-6B). Instead, GLTSCR1 knockout produces a small but significant increase in sensitivity to BRD4 inhibitor (Figure 2-6C, SI 2C). To test the effect of GLTSCR1 on expression of well-characterized BRD4 target, *MYCC*, we measured *MYCC* mRNA levels in the *GLTSCR1* knockout in LNCaP cells and found an increase in MYC levels, which was reversed upon low dose (50 nM) treatment with JQ1 (Figure 2-6D). This provides evidence that GLTSCR1 only slightly modulates BRD4 function in LNCaP cells, potentially by sequestering it from transcriptional activators such as NSD3 that are required for the activation of MYCC transcription by BRD4 [69].
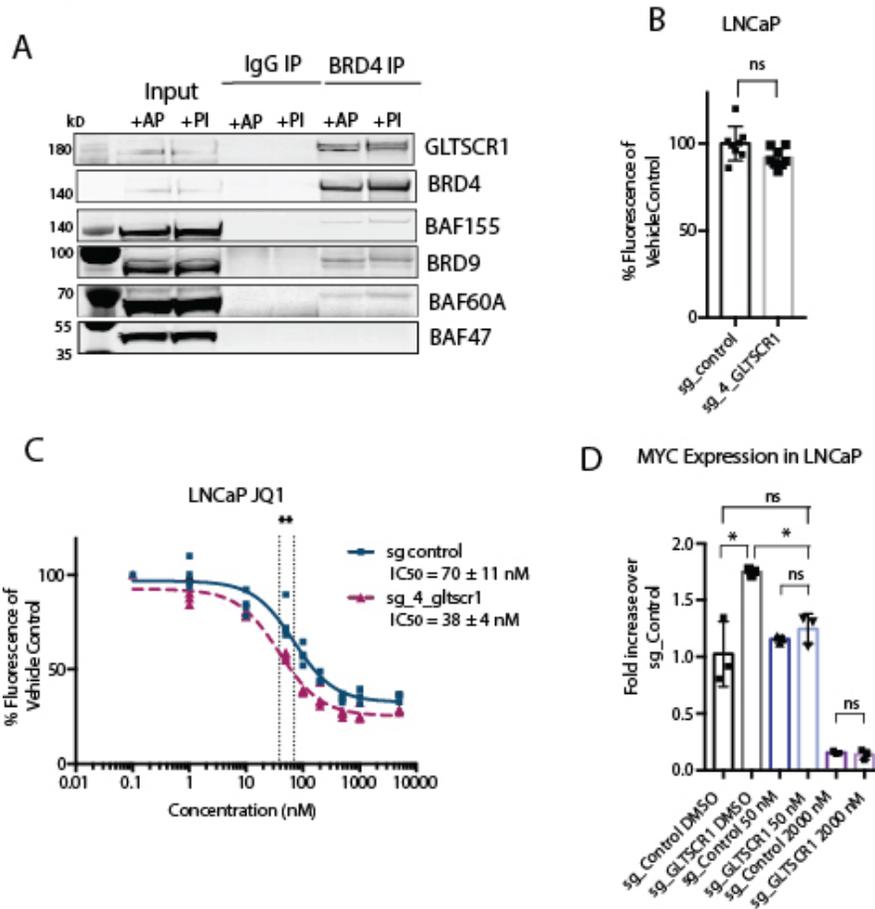
Figure 2-6 GLTSCR1 associates with BRD4, but is not required for BRD4-mediated MYC transcription in LNCaP cells. A) Immunoprecipitation of BRD4 enriches GBAF subunits GLTSCR1, BRD9, BAF155 but not BAF/PBAF subunit BAF47. B) GLTSCR1 knockout did not significantly alter the growth of LNCaP cells. Error bars, mean ± s.d. for n = 6 replicates. C) GLTSCR1 knockout sensitized LNCaP to BET inhibitor JQ1. IC50 values are derived from curve fit calculations using GraphPad Prism and presented as mean ± s.d. for n = 4 replicates. ** p<0.01. D) MYC expression is upregulated in GLTSCR1 knockout LNCaP cells, which reverted back to basal levels upon 50 nM JQ1 treatment. Error bars, mean ± s.d. n = 3 replicates * p<0.05.

We next performed immunoblot analysis to evaluate the expression levels of BICRAL and GLTSCR1 in a series of cell lines. We found that the majority of cell lines have similar expression of these subunits (Figure 2-7A). Gltscr1 knockout (Figure 2-2B) in mouse ESCs did not affect cell viability (Figure 2-7B). In addition, we observed no support for GLTSCR1 as a glioma tumor suppressor as GLTSCR1 knockout in human astrocyte cell line SVG p12 and glioblastoma cell line T98G both resulted in no change in viability (Figure 2-7C). Lastly, we didn't find evidence that BICRAL and GLTSCR1 have redundant, necessary functions as knockout of both GLTSCR1

and BICRAL in HEK293T cells did not produce any viability defect Figure 2-5D, Figure 2-7D). We did, however, detect a dramatic decrease in both proliferation and colony formation upon GLTSCR1 knockout in prostate cancer cell line PC3 (Figure 2-7E, Figure 2-7F). We further knocked out BICRAL in this cell line and found similar defects in cell growth, indicating an overall dependency on GBAF function in this cell line (Figure 2-7E, Figure 2-7F). While PC3 cells are dependent on GLTSCR1 and BICRAL, they have low expression of BRD9 (data not shown), and are not responsive to the BRD9 inhibitor BI-7273 (Figure 2-7G) [33], indicating that GLTSCR1 function is not always dependent on, or synonymous with, BRD9 function.
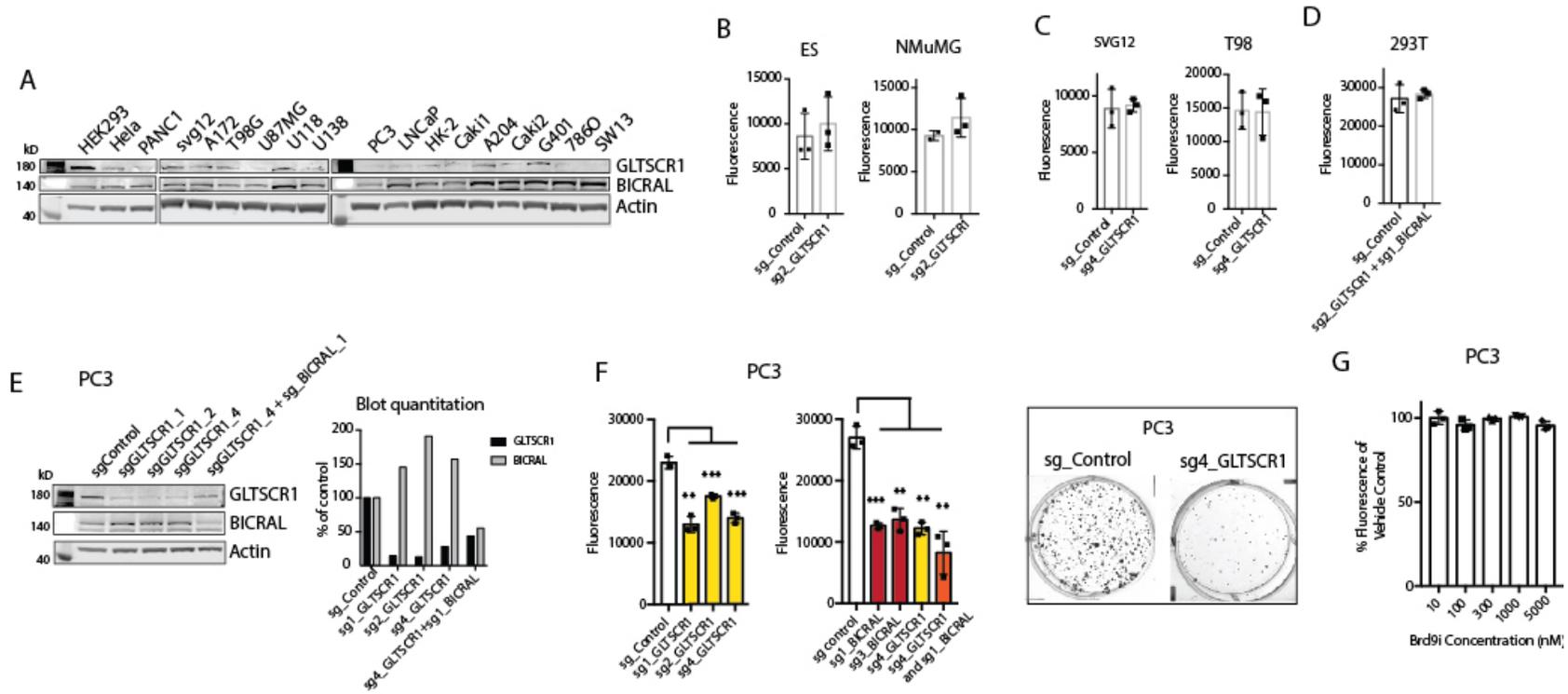
Figure 2-7 GLTSCR1 and BICRAL are expressed in most cell lines but are uniquely required for the viability of prostate cancer cell line PC3. A) Immunoblot analysis of GLTSCR1 and BICRAL expression across a panel of cell lines. B) Proliferation measurement after 6 days of growth of the non-transformed mouse cell lines mESCs and NMuMG with Gltscr1 knockout using Alamar Blue. C) Proliferation measurement after 6 days of growth of the transformed human astrocyte cell line SVGp12 and glioblastoma cell line T98G with GLTSCR1 knockout using Alamar Blue. D) Proliferation measurement after 6 days of growth of HEK293T cells with GLTSCR1 and BICRAL knockout using Alamar Blue. E) Loss of GLTSCR1 reduced the clonogenic growth of prostate cell line PC3. F) (Left) Validation of knockouts using multiple guide RNAs. (Right) Alamar blue assay demonstrated that loss of GLTSCR1 and BICRAL reduced the growth of PC3 cells 6 day after plating. Fluorescence values graphed (excitation 560 nm; emission 590 nm) represent the metric for cell number. Error bars, mean ± s.d. n = 3 biological replicates. ** $p<0.01$ ***$p<0.001$ compared to control cells. G) PC3 cells did not display sensitivity to BRD9 inhibitor BI-7273 (IC50 of 275 nM) up to 10 μM treatment for 4 days. n = 3 biological replicates.

2.3 Discussion

SWI/SNF complexes play diverse roles in normal function and disease; however, most of our understanding of SWI/SNF function is from studying the ATPase subunit BRG1, which is found in multiple different SWI/SNF subcomplexes. The fact that many of the disease-related mutations are in subcomplex-specific subunits has placed increased importance in defining the composition and function of individual SWI/SNF subcomplexes. Our discovery of the ubiquitous new subcomplex GBAF, which is defined by novel subunit paralogs GLTSCR1 and BICRAL, provides another potential mechanism by which BRG1 exerts its functions. We have identified GBAF as a ubiquitously expressed SWI/SNF subcomplex with only a subset of the canonical SWI/SNF subunits but full in vitro ATPase activity. Gene and protein expression data indicate that these paralogs are expressed ubiquitously [63]; however, knockout in many cell lines provides no immediate viability phenotype. Although this complex does not appear to be generally essential for basic cellular viability, mouse knockout data reports an embryonic lethal phenotype for Bicral knockout animals [70]. Whether these developmental roles will be shared with Gltscr1 remains to be seen.

In contrast to the high mutation rates for subunits of the BAF and PBAF complex, subunits of GBAF (with the exception of BRG1) are not highly mutated in cancer [71]. Nevertheless, our data in metastatic prostate cancer cell line PC3 suggests a possible dependency of select cancers on GLTSCR1 and/or BICRAL. Intriguingly, prostate cancers have BRG1 upregulation, but not SNF5, and display dependencies on BRG1 [72]. In addition to prostate cancer, many other cancers display increased dependency on BRG1, although the associated SWI/SNF subcomplex involved in this dependency is unexplored. Inhibitors to SWI/SNF complexes have been proposed as therapies; however, inhibitors of BRG1 ATPase activity will likely have severe toxicity due to the role of BRG1 in general viability in many cell types. Therefore, the development of inhibitors to GBAF subunits may be a more promising approach. Malignant rhabdoid tumors (MRTs) with mutations in SNF5 are dependent on BRG1 [73], and recent reports of MRT sensitivity to BRD9 inhibitors [74] presents the possibility that these cancers might be dependent on GBAF function in some manner. Similarly, AML is dependent on SWI/SNF subunits consistent with GBAF [75], including BRD9, [33,76] providing a potential therapeutic target in these cancers.

Our results also indicate a potential unexplored role for GBAF in BRD4-dependent function. While several studies have noted the association between BRD4 and BRG1 [77], it has not

been clear how they might be functionally related in cancer. For example, AML is dependent on both BRD4 [78,79] and BRG1 [75,79]; however, their roles in AML transcriptional regulation are very different, making it difficult to determine the functional relevance of this association. We find that the association between BRD4 and BRG1 is specific to GBAF, which will provide a framework for deciphering the functional relevance of this association in both the normal and cancer setting. Further defining the importance of the association between GLTSCR1 and BRD4 as well as defining the general contribution of GBAF in chromatin targeting, nucleosome remodeling, and transcriptional regulation will be critical for defining its contribution to human development and disease.

2.4   Experimental procedures

2.4.1   Cell lines and culture conditions

PC3 cells (American Type Culture Collection, Manassas, VA) were grown in F12K (Kaighn's modification) (Corning Mediatech, Inc., Manassas, VA) supplemented with 10% fetal bovine serum (JR Scientific, Inc., Woodland, CA), 100 units/ml penicillin and 100 g/ml streptomycin (Corning Mediatech, Inc., Manassas, VA), 2 mM L-alanyl-L-Glutamine (Corning® glutagro™; Corning Mediatech, Inc., Manassas, VA). HEK293T cells were cultured in DMEM (Corning Mediatech, Inc., Manassas, VA) supplemented with 10% fetal bovine serum (JR Scientific, Inc., Woodland, CA), 100 units/ml penicillin and 100 g/ml streptomycin (Corning Mediatech), 2 mM L-alanyl-L-Glutamine (Corning® glutagro™; Corning Mediatech), 1 mM sodium pyruvate (Corning Mediatech). NMuMG cells were cultured in DMEM (Corning Mediatech) containing 10% fetal bovine serum (JR Scientific, Inc., Woodland, CA), 100 units/ml penicillin and 100 g/ml streptomycin (Corning Mediatech), 2 mM L-alanyl-L-Glutamine (Corning® glutagro™; Corning Mediatech), 1 mM sodium pyruvate (Corning Mediatech) and 10 μg/mL insulin (Sigma). Mouse embryonic stem cell line E14 was cultured in DMEM (Corning Mediatech) supplemented with 15% fetal bovine serum (JR Scientific, Inc., Woodland, CA), 100 units/ml penicillin and 100 g/ml streptomycin (Corning Mediatech), 2 mM L-alanyl-L-Glutamine (Corning® glutagro™; Corning Mediatech), 1 mM sodium pyruvate (Corning Mediatech), 10 mM HEPES (HyClone Laboratories, Inc) , 1% MEM nonessential amino acids (Corning Mediatech), 1x 2-mercaptoethanol (Gibco) and 0.2% leukemia inhibitory factor (LIF)-conditioned media. E14

cells were plated onto gelatinized tissue culture plates without feeder layer. SVG p12 and T98 cell lines were grown in MEM media (Corning Mediatech) supplemented with 10% fetal bovine serum (JR Scientific, Inc., Woodland, CA), 100 units/ml penicillin and 100 g/ml streptomycin (Corning Mediatech), 2 mM L-alanyl-L-Glutamine (Corning® glutagro™; Corning Mediatech), 1 mM sodium pyruvate (Corning Mediatech), 1% MEM nonessential amino acids (Corning Mediatech). LNCaP cells were cultured in RPMI 1640 phenol-free media with 10% fetal bovine serum (JR Scientific, Inc., Woodland, CA), 100 units/ml penicillin and 100 g/ml streptomycin (Corning Mediatech), 2 mM L-alanyl-L-Glutamine (Corning® glutagro™; Corning Mediatech). Caki1 cells were cultured in McCoy's 5A medium (Corning Mediatech) with 10% fetal bovine serum (JR Scientific, Inc., Woodland, CA), 100 units/ml penicillin and 100 g/ml streptomycin (Corning Mediatech), 2 mM L-alanyl-L-Glutamine (Corning® glutagro™; Corning Mediatech). THP1 cells were cultured in RPMI 1640 medium (Corning Mediatech) supplemented with 10% fetal bovine serum (JR Scientific, Inc., Woodland, CA), 100 units/ml penicillin and 100 g/ml streptomycin (Corning Mediatech), 2 mM L-alanyl-L-Glutamine (Corning® glutagro™; Corning Mediatech) and 1x 2-mercaptoethanol (Gibco). All cell lines are incubated in 37°C and 5% CO_2 atmosphere.

2.4.2 Antibodies

Antibodies used in the study are BRG1 (Abcam, ab110641, IP and western blot), BAF60A (Bethyl, A301-594A, IP), BAF170 (Santa Cruz, sc-17838, IP and western blot), GLTSCR1 (Santa Cruz, sc-515086, IP and western blot), FLAG (Sigma Aldrich, F1804), BRD4 (Bethyl, A301-985A50, IP), BRD9 (A303-781A IP and western blot), SS18 (Cell Signaling Technologies, 21792S, IP and western blot), ARID1A (Santa Cruz, sc-32761, western blot), GLTSCR1L/BICRAL (Invitrogen, PA5-56126, western blot), BRD4 (Bethyl, A700-005-T, western blot), BRD7 (Santa Cruz, sc-376180, western blot), ARID2 (Bethyl, A302-230A, western blot), BAF155 (in-house, IP) BAF45D (in-house, IP and western blot), BAF57 (Bethyl, A300-810A, IP and western blot), BAF155 (Santa Cruz, sc-32763, western blot), BAF53A (Abcam, ab131272, western blot), BAF60A (Santa Cruz, sc-514400, western blot), PBRM1 (Bethyl, A301-590A, western blot), ARID1B (Bethyl, A301-047-T, western blot), BCL11A (Santa Cruz, sc-514842, IP and western blot), Actin (Santa Cruz, sc-47778, western blot), GAPDH (Santa Cruz, sc-137179, western blot).

2.4.3 Immunoblot Analysis

Proteins from whole cells, nuclear extracts, salt extractions or glycerol gradient sedimentation analyses were mixed with 4X lithium dodecyl sulfate sample buffer containing 10% 2-merchaptoethanol. The proteins were denatured for 5 minutes at 95 °C, separated on a 4–12% SDS-polyacrylamide gel, and transferred to a PVDF membrane (Immobilon FL, EMD Millipore, Billerica, MA). The membrane was blocked with 5% bovine serum albumin (VWR, Batavia, IL) in PBS containing 0.1% Tween-20 (PBST) for 30 mins at room temperature and then incubated in primary antibodies overnight at 4°C. The primary antibodies were detected by incubating the membranes in goat-anti-rabbit or goat-anti-mouse secondary antibodies (LI-COR Biotechnology, Lincoln, NE) conjugated to IRDye 800CW or IRDye 680 respectively for 1 h at room temperature, and the signals were visualized using Odyssey Clx imager (LI-COR Biotechnology, Lincoln, NE).

2.4.4 Immunoprecipitation

Cells were harvested by trypsinization and washed once in ice-cold phosphate buffered saline (pH 7.2). The pellet was resuspended in Buffer A (20 mM HEPES pH 7.9, 25 mM KCl, 10% glycerol, 0.1% NP-40 with PMSF, aprotinin, leupeptin and pepstatin) at a concentration of 20 million cells per ml. Cells were kept on ice for 5 minutes and nuclei were isolated by centrifugation at 600 g (Eppendorf Centrifuge 5810 R, Hamburg, Germany) for 10 minutes. Pelleted nuclei were washed once in buffer A without NP-40 and pelleted again. The nuclei pellet was resuspended in chromatin IP buffer (20 mM HEPES pH 7.9, 150 mM NaCl, 1% triton X-100, 7.5 mM MgCl2, 0.1 mM CaCl2). 4U/ml Turbo DNase (Ambion, Inc, Foster City, CA) was added to extracts and rotated at 4°C for 30 minutes. The extracts were cleared by centrifugation (Eppendorf Centrifuge 5424 R, Hamburg, Germany) at 21 000 g for 30 minutes. The cleared extract was precleared with normal IgG (Santa Cruz, Dallas, TX) -conjugated Protein A/G magnetic beads (Pierce, Rockford, IL). One microgram specific IgG was used per 0.2 mg lysate for immunoprecipitation. After overnight incubation, immunocomplexes were captured using Protein A/G magnetic beads following 2-hour incubation. The beads were washed twice in chromatin IP buffer and 3 times in high stringency wash buffer (20 mM HEPES pH 7.9, 500 mM NaCl, 1% Triton X-100, 0.5% sodium deoxycholate, 1 mM EDTA). The proteins were eluted in 1X lithium dodecyl sulfate (LDS) loading dye (Thermo Scientific) by boiling at 70°C for 10 minutes. For urea denaturation followed by BRG1 IP, urea

was added into nuclear lysates to final concentration of 0.5 M or 2.5 M and incubated at 4°C for 1 hour. The lysates were then dialyzed against chromatin IP buffer for 50 minutes, precleared and incubated with normal IgG or BRG1 antibodies. For on-bead alkaline phosphatase treatment during BRD4 IP, proteins were extracted in buffers with or without 1X phosphatase inhibitor cocktail 3 (Apexbio, Taiwan) / 1 mM sodium orthovanadate and immunoprecipitated as described above. Following two washes in chromatin IP buffer, beads were washed once in FastAP reaction buffer (Thermo Scientific, Waltham, MA) and incubated at 37°C for 1 hours with or without 10 U alkaline phosphatase. Reaction mixtures were removed, and beads were washed in chromatin IP buffer twice more. Beads were then boiled and run on gel.

2.4.5 Glycerol gradient sedimentation analysis

Thirty million cells were collected by trypsinization; lysed in buffer A and washed once with buffer A without NP-40. Nuclei were resuspended in Buffer C (10 mM HEPES, pH 7.6, 3 mM MgCl2, 100 mM KCl, 0.1 mM EDTA, 10% glycerol). 0.3 M ammonium sulfate was added on nuclei suspension and rotated at 4°C for 30 minutes. Chromatin pellet was removed by ultracentrifugation at 150 000 g for 30 minutes. 0.3 g/ml ammonium sulfate powder was added and the supernatant was incubated on ice for 20 minutes. Proteins were precipitated by ultracentrifugation at 150 000 g for 30 minutes. The protein pellet was resuspended in 100 μL HEMG1000 buffer (25 mM HEPES pH 7.6, 0.1 mM EDTA, 12.5 mM MgCl2, 100 mM KCl) with protease inhibitors. 10-30% glycerol gradient was prepared using HEMG1000 buffer without glycerol and HEMG1000 buffer with 30% glycerol. Resuspended protein was layered over the top of 10-30% glycerol gradient (10 mL) and was fractionated by centrifugation at 40 000 rpm (Beckman Coulter XL-100K, Brea, CA) for 16 hours using SW32Ti rotor (Beckman Coulter, Brea, CA). Twenty 500 μL fractions were collected sequentially from the top and used for immunoblot analysis.

2.4.6 RT-qPCR

RNA was extracted using Trizol (Ambion, Inc, Foster City, CA). cDNA was synthesized using Verso cDNA synthesis kit (Thermo Scientific) using random hexamers. Specific targets were amplified using SYBR Green Master Mix (Roche, USA) and qPCR primers listed: BICRAL

forward 5'- GTTGCCACTCAGCTCCTAAA-3'; BICRAL reverse 5'-CCTCCTGGTTGAACATCCTATC-3'; GLTSCR1 forward 5'-GATGAGGATGGGAGATGCTTAC-3'; GLTSCR1 reverse 5'-TCATAGAAGGCACTTTGGGC-3'; BRG1 forward 5'- TACAAGGACAGCAGCAGTGG-3'; BRG1 reverse 5'- TAGTACTCGGGCAGCTCCTT-3'; BRD9 forward 5'-GCCACGACTCCAGTTACTATG-3'; BRD9 reverse 5'- TCTCCTTCTCGGACTTCTTCT-3'; MYCC forward 5'- AATGAAAAGGCCCCCAAGGTAGTTATCC-3'; MYCC reverse 5'-GTCGTTTCCGCAACAAGTCCTCTTC-3'.

2.4.7 Serial salt extraction assay:

Serial salt extraction assay was performed as published with some modifications [80]. Briefly, 5 million HEK293T cells were harvested by trypsinization and washed once with ice-cold PBS. Cells were lysed in modified Buffer A (60 mM Tris, 60 mM KCl, 1mM EDTA, 0.3M sucrose, 0.5% NP-40, 1mM DTT) with protease inhibitor and nuclei were pelleted. Nuclei were then incubated in 200 μl extraction buffer 0 (50 mM HEPES, pH 7.8, 0.3 M sucrose, 1 mM EGTA, 0.1% Triton X-100, 1 mM DTT, protease inhibitors plus 500 nM JQ1 or DMSO) for 10 minutes, centrifuged at 7000 g for 5 minutes and supernatant was collected as "0 mM fraction". The pellet was then resuspended in 200 μl extraction buffer 100 (50 mM HEPES, pH 7.8, 0.3 M sucrose, 1 mM EGTA, 0.1% Triton X-100, 1 mM DTT, protease inhibitors, 100 mM NaCl, 500 nM JQ1 or DMSO), and processed in the same manner to yield "100 mM fraction". Serial extraction was implemented with extraction buffers containing 200, 300, 400 and 500 mM NaCl. 20 μl-aliquots from each fraction were mixed with 4x LDS loading buffer and run for western blotting.

2.4.8 Growth curve analysis and colony formation assay

For growth curve analysis, 500 or 1000 control or CRISPR-edited cells were plated in 96-well plates. After 6 days, culture medium was refreshed with 1:10 Alamar Blue reagent (Thermo Scientific) and incubated for 3 hours. The fluorescence was measured with excitation at 560 nm and emission at 590 nm using BioTek plate reader. For colony formation assays, 100-200 cells were counted and plated on 6-well plates and allowed to form colonies for 10-15 days. Culture medium was removed and washed twice in ice-cold PBS. Then, cells were fixed in 100% methanol

for 10 minutes at-20°C. Methanol was removed and fixed cells were incubated in 0.5% crystal violet (prepared in 25% methanol) for 10 minutes at room temperature. Excess dye was removed by tap water washes until background was cleared. The images were acquired using ChemiDoc (Biorad, Hercules, CA).

2.4.9 ATPase assay

ATPase assay was performed based on previously published procedure [81] using ADP-GloMax Assay (Promega, Madison, WI) with minor modifications. Twenty-five million (for BRG1 IP) or a hundred million (for GLTSCR1 IP) HEK293T cells were lysed in Buffer A. Pelleted nuclei were extracted for 30 minutes at 4°C using lysis buffer (50 mM Tris, pH 8.0, 150 mM NaCl and 0.2% IPEGAL CA-630, 1 mM DTT, 0.2 mM PMSF and protease inhibitors) at a ratio of 50 million cells per 400 μL buffer. The extract was cleared at 21 000 x g for one hour. One microliter of BRG1 antibody, ten microliters of GLTSCR1 antibody or corresponding amount of normal IgG antibodies were added per 400 μL of cleared extract for overnight immunoprecipitation at 4°C in a rotating wheel. Ten microliters (for BRG1 and rabbit IgG) or twenty-five microliters (for GLTSCR1 and mouse IgG) protein A/G magnetic beads were added to each of 400 μL-IP samples and rotated for two more hours. Beads were washed twice in lysis buffer and then in wash buffer (10 mM Tris pH 7.5, 50 mM NaCl, 5 mM MgCl2, 1 mM DTT and protease inhibitors). The number of beads were adjusted such that material from 25 million (for BRG1 IP) or 100 million (for GLTSCR1 IP) HEK293T cells were included per ATPase reaction. The beads were resuspended in 25 μL-reaction buffer (10 mM Tris pH 7.5, 50 mM NaCl, 5 mM MgCl2, 20% glycerol, 1 mg/ml BSA, 4 mM ATP, 0.5 μg/μL ssDNA, 1 mM DTT and protease inhibitors and incubated at 37°C for one hour on shaker. Beads were separated and the reactions were transferred to 96-well opaque white plate. Twenty-five microliters of ADP-Glo reagent were added per well and gently shaken for one hour at room temperature. Fifty microliters of detection reagent were added per well and further shaken for one hour. Luminescence was detected at 1-second integration time.

2.4.10 Cytotoxicity analysis

Ten thousand (LNCaP) or five thousand (PC3) cells were plated in 100 μl on 96-well plate. Next day, JQ1, OTX015, BI-7372 or DMSO was added and cells were further incubated for 4 days.

Cells were treated with Alamar Blue reagent for 3 more hours and absorbance values were recorded at 570 nm and 600 nm. Percent viability was expressed relative to the DMSO-treated control cells.

2.4.11 Generation of CRISPR/CAS9-mediated knockout cell lines

Short guide RNA (sgRNA) sequences were retrieved from [82] or designed using MIT CRISPR Tool (http://crispr.mit.edu/) or Synthego CRISPR design tool (https://design.synthego.com/) (Table 1). The top and bottom strands of the sgRNA were ordered as single-stranded DNA oligonucleotides from Sigma Aldrich and cloned into lenticrispr v2.0 (a gift from Feng Zhang Addgene plasmid #52961) following the well-established protocol [83]. The vector was packaged into lentivirus using HEK293T cells and the viral particles were concentrated by ultracentrifugation and cell lines were transduced with concentrated virus. Stable lines were generated by puromycin selection. For Gltscr1 KO mouse ESCs, clonal lines were generated.

2.4.12 BICRAL cloning and overexpression

BICRAL ORF was purchased from Novogen, China (cat. no. 762821-2). The ORF was amplified with in-frame C-terminal FLAG tag and 20-bp flanking sequences at both ends with homology to vector using Clontech HiFi PCR premix kit (Takara, USA) and cloned into EcoRI-digested TetO-FUW (a gift from Rudolf Jaenisch Addgene plasmid # 20323) using ligation-free In-Fusion HD cloning kit (Takara, USA). The construct was packaged into lentivirus and delivered into target cells together with pLenti CMV rtTA3 Hygro (w785-1) (a gift from Eric Campeau Addgene plasmid # 26730) for tetracycline inducible expression. Cells were selected with puromycin and hygromycin B. For BICRAL expression, HEK293T cells were treated with doxycycline for six days.

2.5 Contributors

All *in vitro* experiments were performed by Aktan Alpsoy. First mass spectrometry analysis was performed by Dr. Emily Dykhuizen and Dr. Elizabeth Porter.

CHAPTER 3.   CHARACTERIZATION OF DEVELOPMENTAL ROLES OF *GLTSCR1* IN MICE

3.1   Introduction

SWI/SNF complexes have been an attractive venue for cancer studies because of the high incidence of mutant complex subunits in a variety of cancer types. Studies showed that the aberrances in the complexes mostly have causal link with the pathologies.

In addition to pathological conditions such as cancer, SWI/SNF complexes have various direct roles in entire span of embryonic development, stem cell maintenance and lineage differentiation [11,13]. Similar to the observation that aberrations in certain subunits are only critical for specific types of cancers, SWI/SNF subunits are specialized to carry out specific functions in tissue- or lineage- specific manner during development.

3.1.1   Categorization of SWI/SNF complexes in developmental perspective

Previously, three major types of SWI/SNF architecture have been discussed, namely canonical BAF (BAF or cBAF), Polybromo-BAF (PBAF) and GLTSCR1/1L-BAF (GBAF). The combinatorial assembly of monomeric subunits or modules into full complexes is mostly governed by the rules of protein-protein interactions [14]. These subcomplexes coexist in a cell and carry out both redundant and specific functions. However, the major diversification in the complex composition is provided by the paralogous subunits, which make the subcomplexes polymorphic as well.

Considering the number of paralogous subunits that a subcomplex can integrate, the major combinatorial diversity is observed in BAF (cBAF) complex that incorporate BAF45, BAF60, BCL7, BAF53, SS18/CREST, BRG1/BRM, the subunits that have multiple forms (paralogs). Indeed, the alternating expression of these paralogous subunits in tissue- and developmental-stage-specific manner, provides another level of categorization of BAF complex.

Figure 3-1 Developmentally specific SWI/SNF complexes [13]. Mouse SWI/SNF complexes are depicted in the illustration, which may marginally differ from the human counterpart. In pluripotent embryonic stem cells, esBAF is the principle canonical-BAF (cBAF) assembly. esBAF has a homodimer of BAF155 but not BAF170 (while human esBAF has BAF170-BAF155 heterodimer [11]). In addition, it incorporates BAF45D, SS18, ARID1A (BAF250A) and BAF53A. During neuronal differentiation, BRM, BAF170 and ARID1B (BAF250B) are also expressed and incorporates into the complex, forming polymorphic neuronal progenitor BAF (npBAF). Terminal differentiation to postmitotic neurons require further alterations in subunit content, the most significant being substitution of BAF53B in place of BAF53A; CREST in place of SS18 and BAF45B or BAF45C in place of BAF45D [19,54].

Paralog switch does not change the overall architecture of the complex and mostly governed by the relative expression of the paralogous subunits. However, the switches are functionally critical. For example, BAF53B is the defining subunit of the neuronal BAF. It was shown that BAF53B expression and BAF53A suppression are essential for neural development and dendrite formation while BAF53A is critical for neuronal progenitor survival and proliferation [16,84]. Similarly, CREST and BAF45B/C have confined roles in neuronal differentiation and development [54,85]. Paralog-dependent functions have been reported in other lineages as well: For instance, BAF60B but not its paralogs BAF60A or BAF60C was found to be critical for neutrophil and eosinophil formation [86]. BAF60B bridges SWI/SNF complex with CCAAT-enhancer binding protein CEBPε, which is responsible for the gene expression program required for granulocyte differentiation. On the other hand, the other paralog, BAF60C, is essential for heart development

[17,18], where BAF60C enhances the interactions between SWI/SNF complex and cardiac-specific transcription factors such as GATA4, TBX5.

The core SWI/SNF subunits have more pleiotropic effects. For instance, BRG1 and BAF155 are necessary to maintain the pluripotency; however, at the same time, lineage differentiation events are dependent on intact BRG1, suggesting that the core factors adapt different functionalities by altering subunit composition. [11,18,84]. It appears that SWI/SNF functions in different tissues and developmental stages require engagement with the specific set of transcription factors, which is potentially ensured by the specific paralogs, as summarized above. To conclude, we can make an analogy of a Formula 1 racing car where the chassis and engine represent the core SWI/SNF subunits and the tires represent the specific paralogous subunits. To enable safe drive in rain, wet tires need to be put on, while in dry, soft and slick tires are preferred. The car will do the same thing – storm the roads– under any condition provided that correct tires are mounted so that they match with the road/weather conditions (specific TFs).

3.1.2   Current knowledge on the phenotypes associated with subcomplex-specific subunits

SWI/SNF complexes are polymorphic multisubunit complexes. Different tissues and different stages of embryonic development, however, have different level of dependence to each subcomplex or subunit. It is hypothesized that differential ability to associate with distinct set of tissue- or developmental stage-specific factors or the physiological dependence to a subunit/subcomplex-specific functions determines the level of sensitivity to the aberrations in the subunits. It is reported that constitutive knockouts of PBAF-specific subunits *Pbrm1*, *Arid2* and *Brd7* are embryonic lethal, where *Pbrm1* and *Arid2* KO phenotypes are shown to be due to failure in heart development [87–89]. For GBAF, only *Gltscr1l* knockout mice has been reported, which demonstrated embryonic lethality at E9.5 with cardiac hypertrophy (https://www.infrafrontier.eu/search?keyword=EM:01516) [90]. However, we lack this information for *Gltscr1* or *Brd9* to gather a complete picture of GBAF roles during development.

3.1.3   The objective

The overall aim of this study is to understand the developmental roles of *Gltscr1* subunit of GBAF complex and characterize its *in vivo* function by generating *Gltscr1* constitutive knockout mice.

3.2   Approach and preliminary results

3.2.1   Potential roles of *Gltscr1* in stem cell differentiation

We started investigating the possible roles of *Gltscr1* in development with mouse embryonic stem cell (mESC) line E14TG2a. Firstly, we knocked out *Gltscr1* using three different sgRNAs and derived at least 2 different clones per guide RNA (Figure 2-2B). We firstly checked whether *Gltscr1* knockout cells had any change in their growth under undifferentiated conditions (Figure 3-2). It appeared that the growth and colony forming capabilities of E14 cells did not change upon *Gltscr1* knockout.
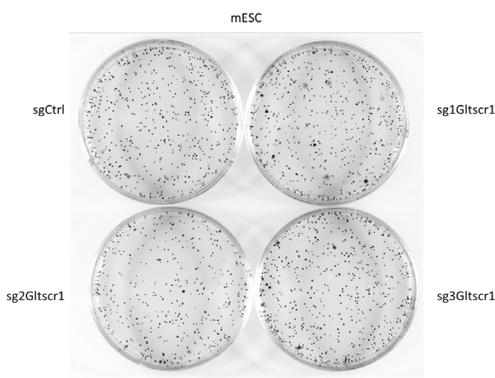
Figure 3-2 Colony formation assay demonstrated that knocking out *Gltscr1* did not change proliferation or colony-forming capacities of the pluripotent (undifferentiated) stem cells.

Next, we tested the effect of *Gltscr1* knockout during differentiation. *In vitro,* mouse embryonic stem cells maintain their pluripotency with the help of a cytokine called leukemia inhibitory factor (LIF). Withdrawal of LIF from the regular culture media leads to spontaneous differentiation of stem cells into all three lineages, visualized as more spread pattern of growth. Morphologically, control cells and *Gltscr1* knockout cells started to spread on culture dish upon withdrawal of LIF, which implies differentiation, without any evident differences in their growth rate. However, tracking the lineage marker genes after initiating the differentiation, we observed

residual but reproducible changes that implied the differences in the direction of the differentiation (Figure 3-3).
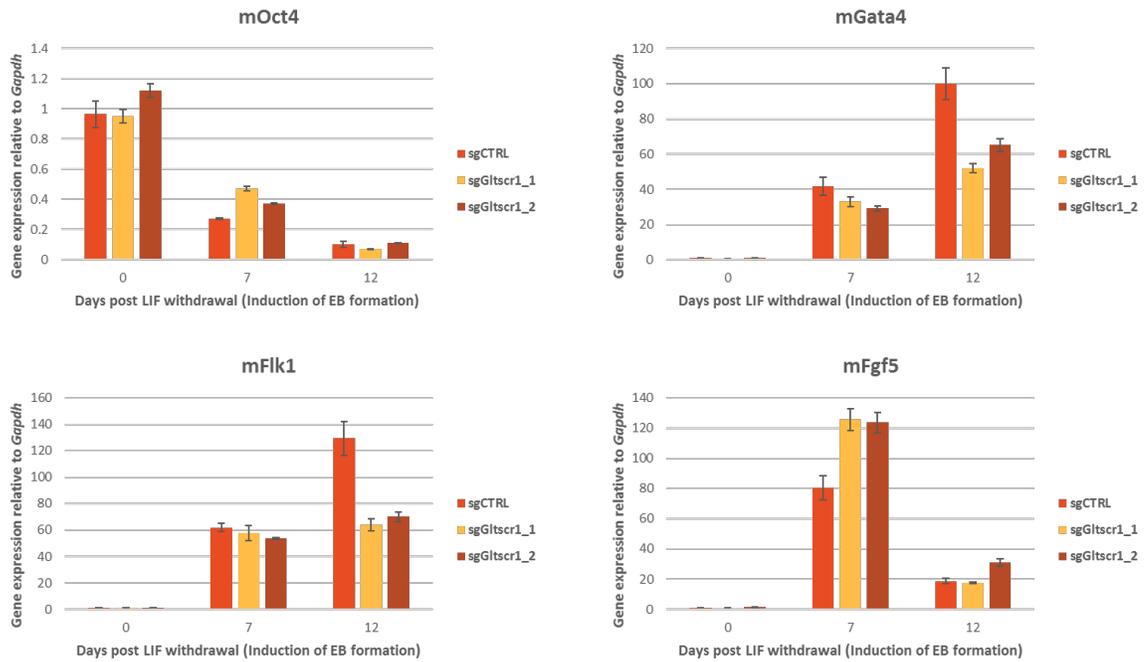
Figure 3-3 *Gltscr1* knockout have altered dynamics of expression in lineage marker genes upon induction of differentiation. Oct4 is the pluripotency factor; Gata4, Flk1 and Fgf5 are endoderm, mesoderm and ectoderm markers, respectively. The same expression profile is obtained with stable *Gltscr1* knockdown mESCs as well (data not shown).

The deviation in the expression of lineage marker genes implied a role for *Gltscr1* during differentiation. Indeed, a recent report suggested that during differentiation via retinoic acid treatment, *Gltscr1* is upregulated while *Gltscr1l* is downregulated in a different mESC line [29], supporting our idea that *Gltscr1* might have roles during differentiation. Because we lack any obvious phenotypic difference in the morphology or growth rate at pluripotent or differentiated states, assessing the role of *Gltscr1* using *in vitro* models of development might be difficult. Next, we decided to generate a constitutive *Gltscr1* knockout mouse, which would enable pinpointing the abnormalities at the whole embryo or organism level.

3.2.2   Generation of constitutive *Gltscr1^{-/-}* mice

Transgenic mice were generated by Purdue University Transgenic Animal Core Facility. Briefly, embryos at blastocyst stage were harvested from c57bl/6 mice and co-injected with mouse

Gltscr1 –targeting small guide RNA (sgRNA) and *Cas9* mRNA. The embryos were transplanted into pseudopregnant c57bl/6 mice. After delivery and weaning, the embryos were genotyped for sequence editing at CRISPR target site by sequencing (Figure 3-4).
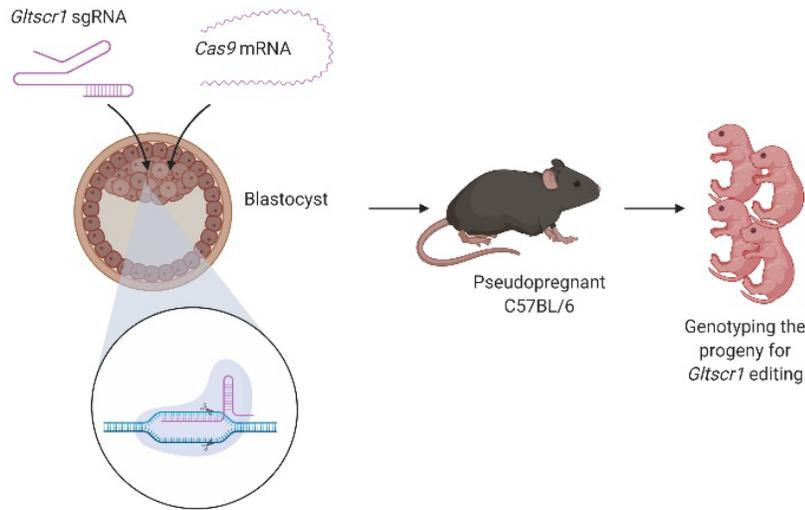
Figure 3-4 Summary of CRISPR/Cas9-mediated generation of *Gltscr1* knockout mice. *Gltscr1* targeting sgRNA and *Cas9* mRNA are coinjected into blastocysts and transplanted into pseudopregnant mice.

Since indels generated by Cas9 without donor plasmid are mostly random, that is, dependent on error-prone non-homologous end joining (NHEJ) repair mechanism, we genotyped the offspring by sequencing a window of 400 base-pair flanking the Cas9 RNP cut site on *Gltscr1* gene. Among 56 young littermates, we detected editing only in 9, all of which contain one edited allele and one wild type allele. Among 9, we had 3 males that bear frameshift deletions (2 with 8-bp deletion and 1 with 10-bp deletion). We then backcrossed them with wildtype females to form F1 generation. Three more backcrosses were performed before setting intercrosses of *Gltscr1*$^{+/-}$ animals to obtain *Gltscr1*$^{-/-}$ mice.

3.2.3 Homozygous mutant mice are perinatal lethal

After three backcrosses, we performed crosses between heterozygous males and females. Surprisingly, upon delivery, we realized obvious differences in some of the pups:

Table 3-1 *Gltscr1^{-/-}* mice is perinatal lethal. The genetic transmission follows Mendelian distribution (Chi-square test; χ=2.552, df=2. Two-tailed p value is 0.2791, suggesting no statistically significant deviation from Mendelian distribution).

Genotype	Born	Postnatal viable
+/+	17	17
+/-	28	27
-/-	22	0
Total	*67*	*44*

Some of the newborns look pale compared to the other pups and have difficulties in breathing, which can be referred as *panting* (Figure 3-5). In our observations, these pups did not survive past 12-hr post-delivery (*postpartum*) and they turned out homozygous mutants. Collectively comparing the genotypes from each litter, we found it fit Mendelian ratios (Table 3-1). Wild type and heterozygous mice survive indistinguishably.

Figure 3-5 *Gltscr1^{-/-}* pups have obvious phenotypic differences with respect to wild type or heterozygous animals. Two pale pups (*left*) are homozygous mutant. They are pale and do not survive more than one day after delivery.

3.2.4 Literature research for the comparable phenotypes

The major observable phenotype of *Gltscr1^{-/-}* mice, that is, the paleness in newborns and embryos has been associated with anemia and erythropoiesis defects [91–93]. Before going into the physiological and molecular details of the defects, mouse erythropoiesis will be briefly discussed in this subsection.

3.2.4.1 Mouse erythropoiesis during development

There are two independent processes of erythropoiesis during embryonic development: Primitive and definitive erythropoiesis. Primitive erythropoiesis starts as early as E7.0 at yolk sac; the immature forms enter circulation as nucleated erythroid precursors and mature in the circulation [94] (Figure 3-6).
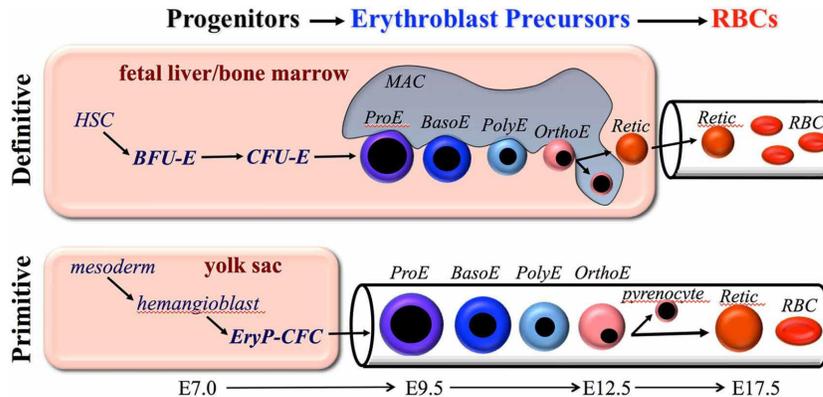
Figure 3-6 Summary of fetal erythropoiesis. Two distinct pathways give rise to circulating RBCs during embryonic development. Primitive erythropoiesis is initiated just after gastrulation. The precursor cells start circulating in the fetal bloodstream and differentiate to RBCs in the circulation. Definitive erythropoiesis, however, starts and progresses in fetal liver. The process is initiated by homing of hematopoietic stem cells (HSCs) in the liver E11.5-E12.5 [95]. The mature, enucleated erythrocytes leave the liver and enter bloodstream E12.5-E16.5, leading to an exponential increase in RBCs in circulation, replacing the primitive erythropoiesis and becoming the key supplier of RBCs for growing embryos. Figure retrieved from [94]

Definitive erythropoiesis starts at fetal liver after homing of HSCs at E11.5-E12.5 [95]. Following the maturation and differentiation in the liver under control of cell-intrinsic and microenvironment factors, this pathway gives rise to exponential burst of mature enucleated RBCs in the bloodstream that replace the circulating primitive erythroid cells (Figure 3-7). It is thought that primitive erythropoiesis supports the embryo's need for the circulating RBCs until definitive erythropoiesis takes over. It was shown that embryo requires erythroid cells in the bloodstream after E9.5 for normal development, i.e., before HSCs colonize in the fetal liver to form the definitive RBCs [95].

Considering the essential roles of *primitive* erythropoiesis in early development, it is expected that aberrations in this process lead to early embryonic lethality [96]. *Gltscr1⁻/⁻ phenotype,*

however, show the characteristics of defects at later stages. To remind, *Gltscr1⁻/⁻* embryos are fully developed and delivered normally–except evident paleness and short survival time after birth. E14.5 embryos were still developing, indistinguishable from normal littermates except mild anemia. Therefore, it is hypothesized that *Gltscr1⁻/⁻*-associated phenotypes are more related with the defects in definitive erythropoiesis (Figure 3-7).
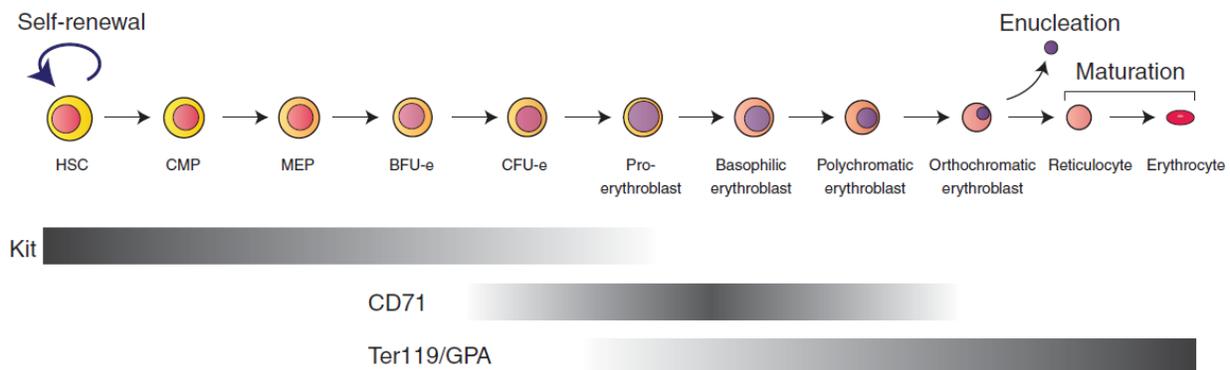
Figure 3-7 Schematic summary of mouse *definitive* erythropoiesis. Multi-step differentiation cascade starting with common progenitor hematopoietic stem cells (HSC) and common myeloid progenitors (CMP) lead to erythroid lineage commitment, firstly forming megakaryocyte-erythroid progenitor (MEP). Further differentiation of MEP forms the definitive erythroid progenitors: Burst-forming unit, erythroid (BFU-e) and more mature progenitors, colony-forming unit, erythroid (CFU-e). Under the effect of multiple soluble factors and receptor-signaling pathways, they give rise to definitive erythroid precursors. Definitive erythroid cells surround a central macrophage in fetal liver, forming erythroblastic islands. This mediates the terminal differentiation and maturation of the definitive red blood cells (RBCs), enucleated erythrocytes [94,97,98]. Stages and identities of the cells are assessed by flow cytometry assays using the specific cell surface markers such as c-kit, CD71 and Ter119. Gradient of grey color represents relative expression levels; light grey for low expression, dark grey for high expression. Figure retrieved from [97]

3.2.4.2 Definitive erythropoiesis defects are generally associated with perinatal lethality

Proper differentiation and maturation of erythroid cells from progenitors involve both intrinsic and extrinsic components. Several signaling pathways, soluble factors and cell-cell contacts are required [91,93,97,99]. Several studies showed that the absence of critical factors such as Gata1 [92]; erythroblast macrophage protein (Emp) [91]; c-Maf [93]; tropomodulin 3 [100] led to more nucleated, immature erythroblasts in peripheral blood and liver preparations of E12.5-E15.5 embryos. In detailed perspective, the mechanism of erythropoiesis failure due to loss of the above factors can be categorized into two main groups: *Cell-autonomous*, that is, the defects which are

reasoned from impaired erythroid progenitors or *non-cell-autonomous*, where the microenvironment that drive the maturation and terminal differentiation of the erythroid cells – such as macrophages in erythroblastic islands are defective. Considering the phenotypic similarities between *Gltscr1$^{-/-}$* and fetal erythropoiesis defects, we tested whether *Gltscr1* is critical for proper fetal erythropoiesis. Indeed, comparing the severity and timing of the defects can give clues about potential failing step in the maturation pathway (Figure 3-5). For instant, defects in the homing of early progenitors or formation of BFU-e and CFU-e have generally been associated with early-stage embryonic lethality.

3.2.5 Preliminary results and discussion

In order to study the erythropoiesis, firstly, timed mating of *Gltscr1$^{+/-}$* intercrosses were set up; the pregnant females at E14.5 stage were euthanized and embryos were harvested. Then, we collected the peripheral blood from embryos and make liver preparations for (1) flow cytometry analysis of erythropoiesis markers Ter119 and CD71; (2) cytospin analysis for blood cells; (3) colony formation assays.

Firstly, roughly one fourth of the embryos per pregnancy looked pale, which are genotyped as *Gltscr1$^{-/-}$*, suggesting that the neonatal anemia is also evident during embryonic development.
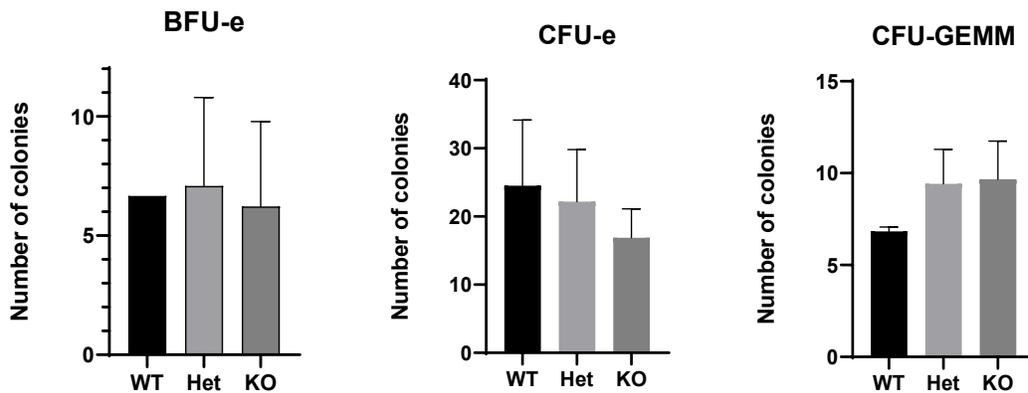
Figure 3-8 Colony forming potential of hematopoietic progenitors from E14.5 fetal livers are not statistically different. Colony types are denoted as burst-forming unit, erythroid (BFU-e); colony forming unit, erythroid (CFU-e) and multipotential CFU-granulocyte, erythrocyte, macrophage, megakaryocyte (CFU-GEMM). Error bars represent standard deviation. n=2, wildtype; n=4 heterozygous; n=3 knockout embryos.

Methylcellulose colony forming assays showed that CFU-e and BFU-e colonies are comparable in number among wildtype, heterozygous and knockout embryos (Figure 3-8). This implied that the knockout phenotype may not be sourced from defects associated with early differentiation of hematopoietic progenitors. To complement this observation, peripheral blood will be collected and analyzed for nucleated: enucleated RBC ratio. Similarly, fetal liver cytospins will be analyzed for nucleated RBCs. We expect to have more nucleated RBCs in KO animals compared to wildtype or heterozygous animals, which has been associated with the anemia [91,93].
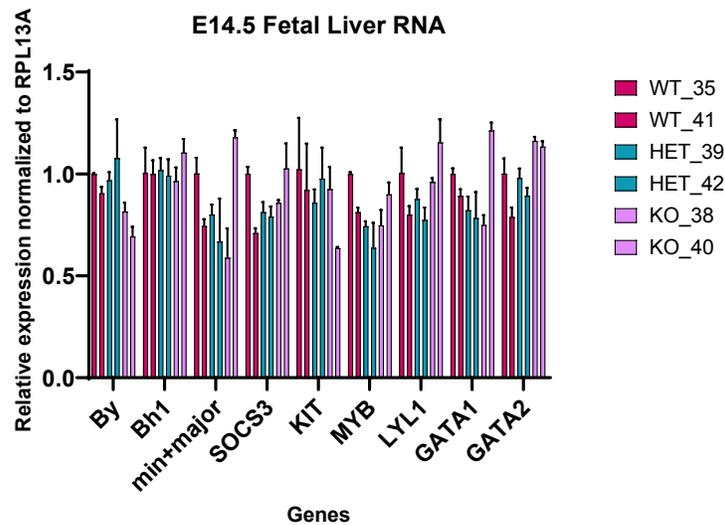
Figure 3-9 RT-qPCR analysis of erythroid cell-autonomous factors from bulk fetal liver cells obtained fresh from wildtype, heterozygous and homozygous knockout animals. Expression of murine hemoglobin genes (embryonic: by, bh1; adult: β^{major}, β^{minor})[101]; transcription factors (Kit, Myb, Lyl1, Gata1/2) were tested. No significant alterations in the expressions was detected.

Whole RNA from E14.5 fetal livers were analyzed for the expression of major cell-autonomous regulators of erythroid differentiation including different classes of globin genes (by, bh1, bminor, bmajor). However, we did not observe a significant difference between wildtype and knockout embryos (Figure 3-9). These suggested that the erythropoiesis failure due to *Gltscr1* knockout may be non-cell autonomous. In order to address this, we are collaborating with Prof. Lessard for blood cell reconstitution experiments in which fetal liver cells from knockout embryos will be implanted into wild type mice whose hematopoietic system is collapsed by radiation. This will answer the question whether *Gltscr1* knockout phenotype is cell-autonomous or not, i.e. whether the knockout phenotype is owing to defects in erythroid precursors or the other components of the fetal liver environment that mediate erythroid maturation. In case of a non-cell

autonomous effect, it is expected *Gltscr1*[-/-] fetal liver cells successfully restore the hematopoietic system including erythrocytes.

As discussed before, one critical effector of erythroblast differentiation is the macrophages, which form erythroblastic islands with immature erythroblasts via juxtracrine and paracrine signaling [91,93]. The central macrophages are required for proliferation, terminal differentiation and enucleation of erythroid cells such that failure to establish macrophage-erythroid contacts results in high nucleated erythrocyte counts in the fetal peripheral blood even though the progenitors retain the capability to form colonies in semisolid media [93].

In parallel to functional assays to address the failing steps in erythroid differentiation, we performed RNA-seq with the bulk fetal liver cells obtained from *Gltscr1*[+/+]*, Gltscr1*[+/-] and *Gltscr1*[-/-] littermates (2 pups for each group) at E14.5. Differentially expressed gene analysis demonstrated that comparing KO with WT pups, 37 genes were upregulated, 193 genes were downregulated (edgeR, FDR cutoff 0.05). Comparing KO with heterozygous pups, 32 genes were upregulated, 177 genes were downregulated. However, compared to wildtype only 53 genes were differentially expressed in heterozygous pups: 29 of them are upregulated and 24 were downregulated. The top GO biological processes terms associated with KO vs. WT comparison turn out to be immune-system related processes such as phagocytosis, cytokine production (Figure 3-10). Similarly, based on "cellular component" GO terms among the most significantly represented terms are ECM or secretory granules. To note, RNA-seq analysis did not demonstrate any significant change in the expression of hemoglobin genes or critical regulators of erythropoiesis (considering n=2 biological replicates), in agreement with RT-qPCR analysis in Figure 3-9.
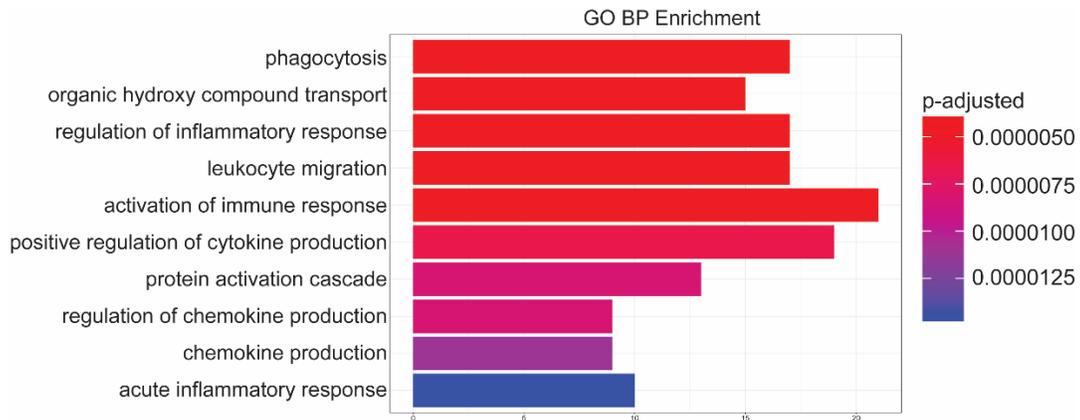
Figure 3-10 GO Biological Processes enrichment plot for the DEG comparison between KO vs. WT. Length of the bars are reflective of number of the genes in GO categories and the color reflects the p-adj values (right).

Considering the requirement of macrophage-erythroblast contacts through cell surface receptors, ECM and possibly soluble factors, our RNA-seq results support the idea of defects in cell-cell contacts and erythroblast homing within erythroblastic islands. Indeed, several differentially expressed genes in *Gltscr1-/-* fetal livers encode cell-surface proteins that are macrophage–dedicated. Those factors include VCAM, leucocyte selectin, colony stimulating factor receptor 1 (Csfr1), Fcgr1 (Table 3-2) [93,102,103]. VCAM is one of the best characterized cell surface receptors of central macrophages in erythroblastic islands [102]. Also, Csfr1 is the master receptor on monocytes and macrophages that transduce the colony-stimulating factor signaling, which is critical for macrophage survival, proliferation and differentiation.

Table 3-2 Relative expression values of the select macrophage signature genes [93,102,103] that are also downregulated in *Gltscr1-/-* fetal livers.

Gene	Fold Change (KO/WT)	p-value
Vcam1	0.682443	4.79E-07
Csf1r	0.581191	3.39E-06
Fcgr1	0.477647	8.17E-05
Sell	0.414804	3.77E-10
Siglec1	0.179532	3.35E-13

Taking together, our RNA-seq data and preliminary functional data suggests that knockout embryos have late-stage defects associated with enucleation, which is possibly due to the defective central macrophage differentiation or macrophage–erythroblast contacts.

3.2.6 Future work

Our future plan includes hematopoiesis reconstitution experiment as discussed, harvesting more peripheral blood and preparing more fetal liver cytospins from different embryonic ages to compare the severities of potential enucleation defects. In addition, we will check whether erythroblastic island forms by immunofluorescence using specific central macrophage (F4/80, VCAM) and erythroblast (Ter119, c-Kit or CD71) markers. We will also isolate the fetal liver associated macrophages and validate our bulk RNA-seq results for the expression of cell surface receptors, cytokine/chemokine and ECM components by RT-qPCR.

In the long term, as a part of mechanistic studies, we are planning to do *ex vivo* analysis of the causal cell types (such as macrophages) for in-dept investigation of Gltscr1 activity.

3.3 Contributors

The embryo injections were performed in Purdue University, Transgenic Animal Core Facility under supervision of Dr. Judy Hallett. RNA-seq data analysis was performed by Dr. Sagar Utturkar. Mouse colonies were maintained in HANSEN Animal Facility; the animal health, bedding and food were routinely monitored by the facility staff. All *in vitro* assays, genotyping and breeding practices were performed by Aktan Alpsoy. Mouse dissection and colony formation assays were performed by Dr. Emily Dykhuizen. Cytospin assays, blood smears and flow cytometry assays were performed by Aktan Alpsoy, Dr. Emily Dykhuizen, Dr. Alisha Dhiman. Dr. Emily Dykhuizen and Dr. Alisha Dhiman analyzed flow cytometry data.

CHAPTER 4.   ELUCIDATING THE ROLES OF GBAF IN PROSTATE CANCER

4.1   Introduction

Prostate cancer is one of the leading causes of cancer-related deaths in men worldwide. Localized and androgen-responsive early-stage prostate cancer is generally treated by surgery, radiotherapy, while the primary regimen for treatment or containment of metastatic tumors is androgen deprivation. Androgen deprivation involves chemical or physical castration that primarily aims to deplete the synthesis of androgens thereby reducing androgen signaling. However, clinically, this approach is prone to fail eventually due to emergence of resistance to low androgen levels, which is termed as castration resistance. Currently, castration resistant prostate cancer (CRPC) is treated with anti-androgens such as enzalutamide, and abiraterone; however, resistance to these drugs still develops.

It has been known that androgen receptor-mediated transcription is dependent on several coregulatory proteins [104]. These proteins or protein complexes have roles in recruitment of AR, making chromatin regions more accessible for AR and enhancement of AR transactivation. In recent years, targeting the coregulatory proteins of androgen receptor instead of or together with androgen receptor have been proposed as novel therapy options to overcome hormone therapy-related failures. One critical advancement has been the elucidation of the role of bromodomain and extra terminal domain (BET) proteins in AR-mediated gene expression [68,105,106]. These studies demonstrated that BET proteins have therapeutic value in prostate cancer; mechanistically it was shown that BET proteins directly interact with AR and they are required for localization of AR on the genome as well as expression of target genes [68,106,107]. Similarly, components of switch/sucrose nonfermentable (SWI/SNF) chromatin remodeler complexes play critical roles in AR-mediated transactivation [108–111]. Although 20% of malignancies have distinct SWI/SNF mutations [19], prostate cancer rarely has SWI/SNF mutations [112], which may imply utility of intact SWI/SNF complexes for prostate cancer. Indeed, a recent study identified BRG1, the ATPase subunit of SWI/SNF complexes, as an essential factor for survival of *PTEN*-null prostate cancer [113].

SWI/SNF complexes assume compositionally and functionally distinct subcomplexes. Previously, we and others defined a novel type of SWI/SNF subcomplex that is termed as GLTSCR1/1L-BAF (GBAF) or noncanonical BAF (ncBAF) that lacks several core BAF subunits

such as ARID, BAF47, BAF57 while uniquely possessing GLTSCR1/1L or BRD9 [27–31]. These studies identified GBAF as specific vulnerability of SWI/SNF-perturbed cancer types such as synovial sarcoma [27,30] and rhabdoid tumors [27,31]. While the roles of core SWI/SNF subunits have been relatively well-studied for prostate cancer, the specific roles of GBAF in SWI/SNF-intact cancers -such as prostate cancer, has not been addressed. In this report, we demonstrated that GBAF complex is critical for maintenance of prostate cancer cells *in vitro* and *in vivo*. Both small molecule inhibitor against non-redundant unique subunit BRD9 or its knockdown altered the AR target gene expression. Mechanistically, we found that BRD9 interacts with AR and BET proteins –specifically BRD2 and BRD4. In addition, BET inhibitor I-BET151, antiandrogen enzalutamide and *BRD9* had significantly overlapping transcriptional effects. We also found co-enrichment of BRD9, AR and BRD4 genome-wide in LNCaP cells. Overall, this study suggested GBAF acts in concert with BET proteins and AR to regulate prostate cancer oncogenesis. Targeting GBAF via its bromodomain-containing subunit BRD9 can serve as an alternative therapy option for the prostate cancer.

4.2 Materials and Methods

4.2.1 Cell lines and cell culture

LNCaP (clone FGC), VCaP, PC3, HEK293T cells were purchased from American Type Culture Collection (ATCC, Manassas, VA). 22Rv1 and C4-2 cells are kind gift from Chang-Deng Hu. HEK293T and VCaP cells were cultured in DMEM media supplemented with 10% fetal bovine serum, 100 units/ml penicillin and 100 g/ml streptomycin, and 2 mM L-alanyl-L-glutamine (Corning Glutagro^{TM}) and 1: 10000 plasmocin (Invivogen, San Diego, CA). LNCaP and 22Rv1 cells were cultured in RPMI 1640 and PC3 cells were cultured in F12K media with the same supplements as above. C4-2 cells are cultured based on ATCC suggestions. All media and supplements were obtained from Corning Mediatech, Inc. For androgen deprivation studies, charcoal-dextran stripped FBS was used at 10% with phenol red-free DMEM or RPMI 1640.

4.2.2 Antibodies

Table 4-1 List of antibodies

Antibody	Company / Catalog number	Application / Dilution
BRD9	Active Motif 61537	WB 1:500; IP 6 µL/mL; ChIP 4 µL/ ~5 million cells
BRD4	Bethyl A301-985A	IP 2 µL/mL; ChIP 2 µL/ ~5 million cells
BRD4	Bethyl A700-004	WB 1:1000
BRG1	Abcam	
BRG1	Santa Cruz sc-17796	WB 1:1000; IP 2 µL/mL
GLTSCR1	Santa Cruz sc-515086	WB 1:200; IP 10 µL/mL
GLTSCR1L	Thermo Scientific PA5-56126	WB 1:500
BRD2	Santa Cruz sc-393720	WB 1:500
BRD2	Bethyl A302-582A	IP 4 µL/mL; ChIP 6 µL/~5 million cells
BRD3	Santa Cruz sc-515666	WB 1:500
AR	Cell Signaling Technologies #5153	ChIP 10 µL/~5 million cells; WB 1:1000
AR	Santa Cruz sc-7305	WB 1:200
BAF47	Santa Cruz sc-166165	WB 1:1000
BAF155	Santa Cruz sc-32763	WB 1:1000
GAPDH	Santa Cruz	
Alpha-tubulin	Santa Cruz	
EZH2	Cell Signaling Technologies	
TBP	Abcam ab818	WB 1:1000
CTCF	Santa Cruz sc-271514	WB 1:500
V5	Cell Signaling Technologies #13202	IP 4 µL/mL
V5	Invitrogen R960-25	1:4000
cPARP	Cell Signaling Technologies #9541	WB 1:500

4.2.3 Compounds

I-BRD9, JQ1, OTX015 were purchased from Cayman Chemicals (Ann Arbor, MI). dBRD9 was purchased from Tocris.

4.2.4 Generation of stable cell lines

GLTSCR1L and *GLTSCR1* knockout LNCaP cell lines were generated as previously discussed [28]. *BRD9*-targeting shRNA sequences, shBRD9#34 (TRCN0000127634) and shBRD9#81 (TRCN0000131081) were cloned into plko.1 puro (Addgene plasmid ID #8453) and tet-plko-puro (Addgene plasmid ID #21915). tet-plko-puro backbone was used for inducible

knockdown constructs. Lentiviral particles were produced by cotransfecting shRNA vector into HEK293T with pHR'-CMV-8.2ΔVPR packaging and pHR'-CMV-VSVG envelope vectors, gifted by Chang-Deng Hu. MSCV-BRD9_N216A-PGK-Puro-IRES-GFP was a gift from Christopher Vakoc (Addgene plasmid # 75116 ; http://n2t.net/addgene:75116 ; RRID:Addgene_75116). MSCV-BRD9-PGK-Puro-IRES-GFP was a gift from Christopher Vakoc (Addgene plasmid # 75114 ; http://n2t.net/addgene:75114 ; RRID:Addgene_75114). Wild-type BRD9 and bromodomain-mutant BRD9 (BRD9-N216A) are amplified with N-terminal V5 tag and subcloned into EcoRI-digested TetO-FUW (a gift from Rudolf Jaenisch Addgene plasmid # 20323) using ligation-free In-Fusion HD cloning kit (Takara, USA).

4.2.5 Cell proliferation and survival assays

LNCaP, VCaP and 22Rv1 cells expressing *BRD9*-targeting hairpins were plated at density of 300 000/well into a 6-well plate 6-days post-transduction. *BRD9*-knockdown HEK293T cells are plated at a density of 5000 cells or 20 000 cells/well. The plates were fixed in ice-cold methanol and stained with crystal violet to assess the viable cells.

LNCaP, VCaP, 22Rv1, C4-2 cells were plated on white 96 well plates and treated with I-BRD9 or dBRD9 for 4-5 days in complete media. Percent viability was assessed using Cell Titer Glo reagent (Promega, Madison, WI). For exogenous GLTSCR1L and BRD9 expression, 3000-4000 cells were plated on 96-well plate with 200 ng/mL or 20 ng/mL doxycycline, respectively. Then the cells are treated with enzalutamide for 4-5 days. Viability was measured using Cell Titer Glo reagent.

4.2.6 RT-qPCR assays

RNA was extracted using TRIzol (Ambion, Inc.). cDNA was synthesized using Verso cDNA synthesis kit (Thermo Scientific) using blend of oligo dT: random hexamer oligonucleotides. Specific targets were amplified using PowerUp SYBR Green Master Mix (Applied Biosystems, Foster City, CA) in Biorad CFX qPCR instrument. For I-BRD9 inhibition studies, 200 000-300 000 cells were plated into a 6-well and treated with various concentrations of the inhibitor for 4-5 days, then RNA was harvested. For the induction studies, cells were plated in charcoal-stripped FBS containing medium. Next day, the cells were pretreated with 5 μM (for

VCaP) or 10 μM (for LNCaP) I-BRD9 for 24 hours before inducing AR target genes by switching to regular medium. RNA was harvested 7-8-hour and 24-30-hour post-induction. For GLTSCR1L or BRD9 expression studies, the cells were plated in charcoal-stripped FBS containing medium and the gene expression was induced with 200 ng/mL (for GLTSCR1L) or 20 ng/mL (for BRD9) doxycycline. Forty-eight hours post-induction, the media was replaced with regular medium and RNA was harvested 8 hours or 24 hours later.

4.2.7 Immunobloting

Whole cell lysates or nuclear lysates were quantified using BCA Assay (Pierce, Rockford, IL). Equal amounts of protein were electrophoresed on Novex 4-12% precast gel or 10% hand-cast gels; transferred onto 0.45μm PVDF membrane and blocked 1 hour in 5% BSA in TBST. After overnight incubation in primary antibodies, the membrane was incubated with infrared-dye labeled Licor goat anti-mouse or anti-rabbit antibodies. The images were acquired using Licor imaging system.

4.2.8 Preparation of samples for ChIP-qPCR and ChIP-seq

Eight million LNCaP cells were plated on 15 cm dishes in regular media. Two days later, the cells treated with 10 μM I-BRD9 for 24 hours, which yield 80% confluence and 15-20 million cells/plate at the time of harvest. The plate was rinsed in PBS + 1 mM $MgCl_2$ three times and crosslinked in 1% formaldehyde for 10 minutes at RT with agitation. Crosslinking was quenched with 125 mM Glycine and the plate washed three times with ice-cold PBS. The cells were scrapped into PBS and pelleted at cold. The pellet was flash-frozen and stored at -80 until the time of processing. The pellet was thawed on ice and sequentially extracted and washed in L1, L2 and L3 buffers with protease inhibitors [114]. After the final wash, the pellet was resuspended in sonication buffer 10 mM Tris pH 8.0; 1 mM EDTA; 0.1% SDS and sonicated using Branson 250 probe sonicator at 20% amplitude with 0.5s on and 1.5s off cycles for 8 minutes (2 minutes on). The sonicated chromatin was cleared by centrifuging at 21 000 x g, for 30 minutes. A 50 μL-aliquot was separated for testing the shearing efficiency and determination of DNA concentration. The chromatin was diluted 1:1 with 2x ChIP dilution buffer (100 mM HEPES pH 7.5; 600 mM NaCl, 1 mM EDTA, 2% Triton X-100, 0.2% sodium deoxycholate, 0.1% SDS) and protein amount was

estimated using BCA assay. Chromatin equivalent of ~5 million cells was incubated with primary antibodies overnight and then incubated with Dynabead A for 2 hours. Then the beads were washed sequentially in low-salt buffer (20mM HEPES pH 7.5, 0.1% SDS, 0.1% Deoxycholate, 1% Triton, 150mM NaCl, 1mM EDTA, 0.5mM EGTA); twice in high-salt buffer (20mM HEPES pH 7.5, 0.1% SDS, 0.1% Deoxycholate, 1% Triton, 500 mM NaCl, 1mM EDTA, 0.5mM EGTA); LiCl wash buffer (20mM HEPES pH 7.5; 0.5% Deoxycholate, 0.5% IPEGAL CA-630; 250 mM LiCl, 1mM EDTA, 0.5mM EGTA); final wash buffer (20mM HEPES pH 7.5, 1mM EDTA, 0.5mM EGTA). The captured chromatin was eluted once with 200 μL and once with 100 μL elution buffer (100 mM $NaHCO_3$, 1% SDS) for 30 minutes at 30°C with shaking. The eluate was supplemented with 12 μL 1 M Tris pH 6.6, 6 μL 0.5 M EDTA, 18 μL 5 M NaCl and 2 μL RNase A and incubated for 30 minutes at 37°C. 2 μL Proteinase K was added and incubated for 16 hours at 65°C. DNA was extracted once with phenol: chloroform; once with chloroform. DNA was precipitated by adding 1/10 volume of 3 M NaOAc pH 5.2; 1 volume 2-propanol and 2 μL glycogen and incubating overnight at -20°C. After centrifugation at top speed for 1 hour, the pellet was washed with 1 mL fresh 70% ethanol and then air-dried. DNA was resuspended in low-EDTA TE and directly used in qPCR. For ChIP-seq, DNA quality and concentration were determined using Agilent Bioanalyzer and FluorNanodrop by Purdue Genomics Core Facility. The samples around 200-400 bp average size were proceeded with library preparation using Ovation Ultralow System V2 UDI according to manufacturer's directions. Quality of libraries was checked and the libraries were submitted for sequencing in NovaSeq 6000 platform (Novogene, Sacramento, CA).

4.2.9   Sample preparation for RNA-seq and analysis

LNCaP cells were transduced with lentiviruses encoding scrambled shRNA, *BRD9* shRNA (shBRD9_81), control guide RNA (sgCtrl), *GLTSCR1* sgRNA (sgGLTSCR1) or *GLTSCR1L* sgRNA (sgGLTSCR1L) and incubated for 2 days. The cells were selected for 6 days with puromycin and plated into 6-well plates in triplicates. Eleven-day post-transduction the cells were harvested in Trizol and purified. *GLTSCR1* knockout set was sequenced by Purdue Genomics Core Facility; *BRD9* knockdown and *GLTSCR1L* knockout sets were sequenced by Novogene. Differentially expressed genes were calculated using edgeR using padj<0.5 and hg19 reference genome. We used publicly available RNA-seq data for I-BET151 treatments (0.5 μM I-BET151; GEO GSE103907) and enzalutamide treatments in LNCaP cells (GEO GSE110903).

4.2.10 Mouse xenograft studies

Five-weeks old male NCr nude mice (Taconic, Rensselaer, NY) were subcutaneously injected from the right flank with 3 million LNCaP tet-on shBRD9 cells in 200 µL PBS: Matrigel (1:1) mixture. Following injection, the mice were randomized to receive regular drinking water or 1.2mg/mL doxycycline hyclate in drinking water. The water was refreshed every two days and tumor size was measured using a caliper. The same cell line was injected into 14 Nrg mice at the same density. After the tumor size reached 200 mm^3, the mice were randomized, and doxycycline treatment started.

4.2.11 Immunoprecipitations

Forty million LNCaP or VCaP cells were harvested by trypsinization following relevant treatments. The cells were washed in ice-cold PBS and lysed in Buffer A (20 mM HEPES, pH 7.9, 25 mM KCl, 10% glycerol, 0.1% Nonidet P-40 with PMSF, aprotinin, leupeptin, and pepstatin) at a concentration of 20 million cells/ml. After 7-minute-incubation on ice, nuclei were pelleted at 600 x g for 10 minutes and nuclei were resuspended in lysis buffer (20 mM HEPES, pH 7.5, 150 mM KOAc, 0.2% NP-40, 1 mM MgCl$_2$) containing protease inhibitors and benzonase at 100U/ml. For the IPs with inhibitor treatment, JQ1 or I-BRD9 was spiked in at 20 µM. The lysis proceeded at 25°C for 40 minutes on a thermoshaker. The lysate was cleared at 21 000 x g, for 30 minutes at 4°C. Cleared lysate was quantitated using Pierce 660nm protein assay. Equal amounts of lysates (200-300 µg per IP) were incubated with Protein A Dynabeads preconjugated to rabbit IgG, BRD4, BRG1, BAF155 or BRD9 antibodies in a rotator at 4°C for 3 hours. The beads were washed 3 times in wash buffer (20 mM HEPES pH 7.5, 100 mM KCl, 10% glycerol and 0.1% NP-40, with protease inhibitors). Beads were resuspended 1x LDS sample buffer and proteins were eluted by shaking at 90 °C for 10 minutes. Immunoprecipitates and inputs (5%) were loaded on 4-12% polyacrylamide gels and separated for western blotting.

4.2.12 Primers used in the study

Table 4-2 List of primers

RT-qPCR	Forward primer	Reverse primer
TMPRSS2	CAGGAGTGTACGGGAATGTGATGGT	GATTAGCCGTCTGCCCTCATTTGT
FKBP5	TCTCATGTCTCCCCAGTTCC	TTCTGGCTTTCACGTCTGTG
MAF	CTGGCAATGAGCAACTCCGA	AGCCGGTCATCCAGTAGTAGT
MAK	AACCGATACACAACCATGAGAC	CCGGATTCATTACTCTTGCCC
CENPN	TGAACTGACAACAATCCTGAAGG	CTTGCACGCTTTTCCTCACAC
AR	CAGTGGATGGGCTGAAAAAT	GGAGCTTGGTGAGCTGGTAG
ARv7	CAGGGATGACTCTGGGAGAA	GCCCTCTAGAGCCCTCATTT
TMEFF2	GACACTGTGACTTGCGTCTGT	CAGGTAACACTCATTCTGGTAGC
SYT4	ATGGGATACCCTACACCCAAAT	TCCCGAGAGAGGAATTAGAACTT
PSA	AGTGCGAGAAGCATTCCCAAC	CCAGCAAGATCACGCTTTTGTT
BCHE	GTCAGAGGGATGAACTTGACAG	TGAATCGAAGTCTACCAAGAGGT
KLK2	TCAGAGCCTGCCAAGATCAC	CACAAGTGTCTTTACCACCTGT
BMPR1B	CTTTTGCGAAGTGCAGGAAAAT	TGTTGACTGAGTCTTCTGGACAA
NSE	CCGGGAACTCAGACCTCATC	CTCTGCACCTAGTCGCATGG
TUBB3	ATCAGCAAGGTGCGTGAGGAGTAT	TCGTTGTCGATGCAGTAGGTCTCA
ChIP-qPCR		
tmprss2-upstream ctcf_1	GAGTAAGGCAGGGTGGATCC	GGGACGTCTTGTAGTGGAGA
tmprss2-upstream ctcf_2	GCTGTTCCCGGACAACTTTT	AGCAGCCCGATGATTGAGG
fkbp5-upstream ctcf	ACCAGTACGTGTGACTACCG	CGCGATAATGCTTTGCTCCT
elovl7-upstream ctcf	ATCTACTGCCGAGAACCCAG	CTTGTCAGTCTTCGCTTCCG
nkx3-upstream ctcf	CTTCTCCTTTGCTCCTCCCA	CGTGGGCAAGAAAAGACACA
TMPRSS2 intron	AACAAGGAAGCCAGGGAAGG	CTCTGGGTCCCATGGCTTTT
BMPR1B enhancer	AACCCCGGAAGCTTTCTCTATC	TGGATTCCTCTGCTCTCCGTT
KLK3 enhancer	TGTCTGCAGGACAGTCTCAAC	GCTGCCAGACACAGTCGAT

4.3 Results

4.3.1 Depletion of *BRD9* reduced the viability of prostate cancer cells

Previously, we demonstrated that depletion of GBAF-specific subunit *GLTSCR1* did not drastically alter the viability of normal or transformed cell lines representative of different cancer

types. However, androgen-negative prostate cancer cell line PC3 displayed a slower proliferation rate upon *GLTSCR1* and *GLTSCR1L* knockouts [28]. In addition, TCGA prostate cancer dataset (Figure 4-1) as well as tissue microarray study [115] revealed that prostate cancer tissues have higher GLTSCR1 expression than normal prostate tissue. In addition, higher GLTSCR1 expression has been associated with certain clinical parameters such as tumor invasion and metastasis [115].
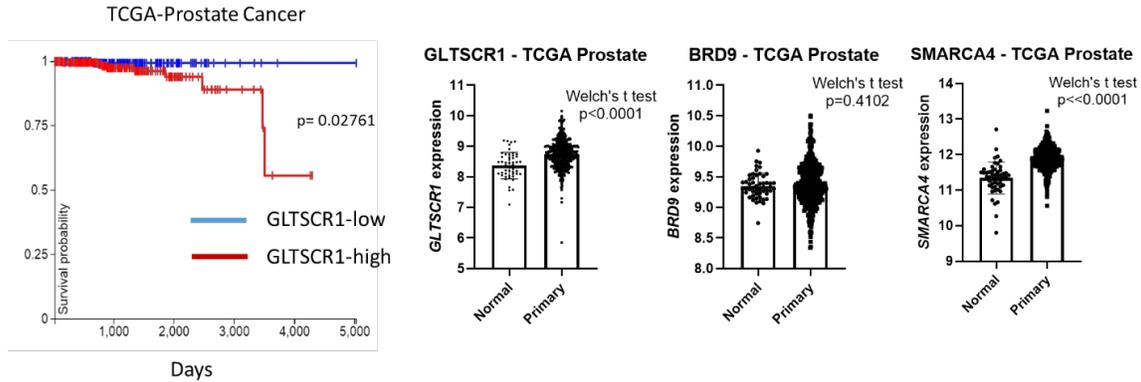
Figure 4-1 GLTSCR1 expression is associated with poor survival in prostate cancer. Kaplan Meier plot implied that patients with high-GLTSCR1 expression has reduced survival compared to patients with low GLTSCR1 expression. Significantly higher RNA levels of *GLTSCR1* and *SMARCA4 (BRG1)* have been detected in clinical samples.

In order to gain more insight for a potential GBAF involvement in prostate cancer, we knocked out individual paralogs *GLTSCR1* and *GLTSCR1L* in another prostate cancer cell line, LNCaP and observed that despite transient retardation in the growth, knockout of both proteins have similar growth profiles (Figure 4-2).
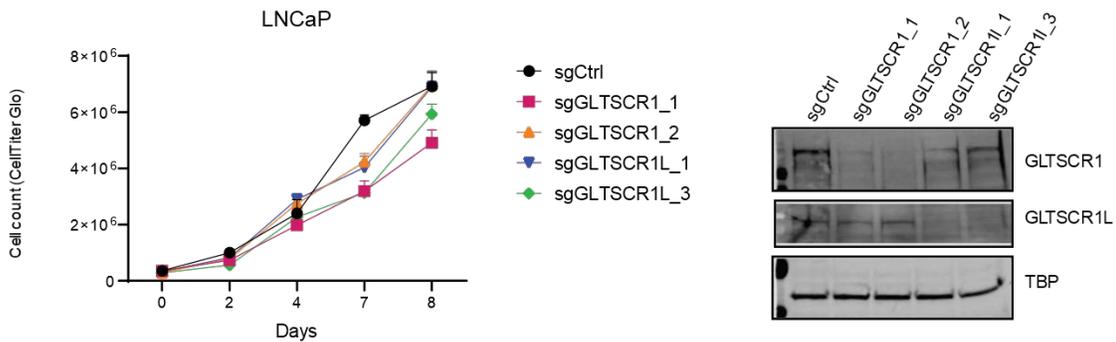
Figure 4-2 GLTSCR1 and GLTSCR1L have mild effect on prostate cancer cells. *GLTSCR1* or *GLTSCR1L* has slight effects on growth of LNCaP cells. Viability of cells that had been plated at equal density into 96 well plates was measured every other day over 8-day period using CellTiter Glo. The luminescence values were normalized to the mean luminescence from day 0 control cells (sgCtrl). Western blot showing the efficiency of *GLTSCR1* and *GLTSCR1L* knockouts.

70

Previously, it was shown that GLTSCR1 and GLTSCR1L are two mutually-exclusive paralogous subunits in GBAF complex, which can biochemically substitute one another [28]. Nevertheless, both being a non-redundant subunit unique to GBAF complex and containing a "druggable" bromodomain, we hypothesized that BRD9 is more therapeutically relevant subunit to probe the function of GBAF. To this end, we knocked down *BRD9* in various prostate cancer cell lines (LNCaP, VCaP and 22Rv1) using lentiviruses encoding either control shRNA (shScramble) or two different *BRD9*-targeting shRNA sequences (shBRD9_34 and shBRD9_81) and validated the knockdown using western blot **(**Figure 4-3a**)**. After transduction and selection using puromycin, we plated the cells on 6-wells for viability test using crystal violet. We observed that androgen-receptor positive cell lines LNCaP, VCaP and 22Rv1 were highly sensitive to loss of *BRD9* (Figure 4-3a). As a bromodomain-containing protein, BRD9 is a druggable protein. In addition to knockdown approach, we tested the efficacy of BRD9-specific bromodomain probe I-BRD9 on the same cell lines. Four days treatment with various concentration of the compound led to a dose-dependent reduction in cell viability (Figure 4-3b). As a tertiary approach, we used a recently-developed degrader against BRD9 that utilizes a different bromodomain probe than I-BRD9, called dBRD9. The degrader at 0.5 μM concentration caused reduced proliferation in 6 days in the same set of prostate cancer lines (Appendix A Supplementary Figure 1b). We also observed that androgen receptor (AR)–negative prostate cancer cell line PC3 was retarded by I-BRD9 and *BRD9* knockdown, a defect that is milder than the AR-positive cell lines (Appendix A Supplementary Figure 6). BRD9 dependence of AR-expressing prostate cancer cells was further validated using embryonic kidney cells, HEK293T and immortalized normal prostate cell line RWPE-1, both of which had indifferent viabilities between control cells and *BRD9* knockdown cells 11-15 days post-transduction (Appendix A Supplementary Figure 1c,d). Besides cell culture models, we also investigated the effect of *BRD9* depletion *in vivo*. To this end, we generated doxycycline-inducible *BRD9*-shRNA expressing LNCaP line and injected into NCr nude mice subcutaneously. Starting the doxycycline treatment immediately after injection, we noticed that doxycycline group had smaller tumors compared to control groups (Figure 4-3c). In order to see whether *BRD9* knockdown would be as effective after tumor formation, we injected the same cell lines into NRG mice; allowed tumors to reach average size of 200 mm^3 before starting the doxycycline-mediated knockdown. Similarly, we observed that further tumor growth was blocked in knockdown group, which had overall significantly smaller tumor mass compared to control

group (Figure 4-3c). However, induction of *BRD9* knockdown in PC3 cells did not affect the tumor growth (Appendix A Supplementary Figure 6), suggesting that the mild effect in PC3 2D growth does not translate to xenograft models.
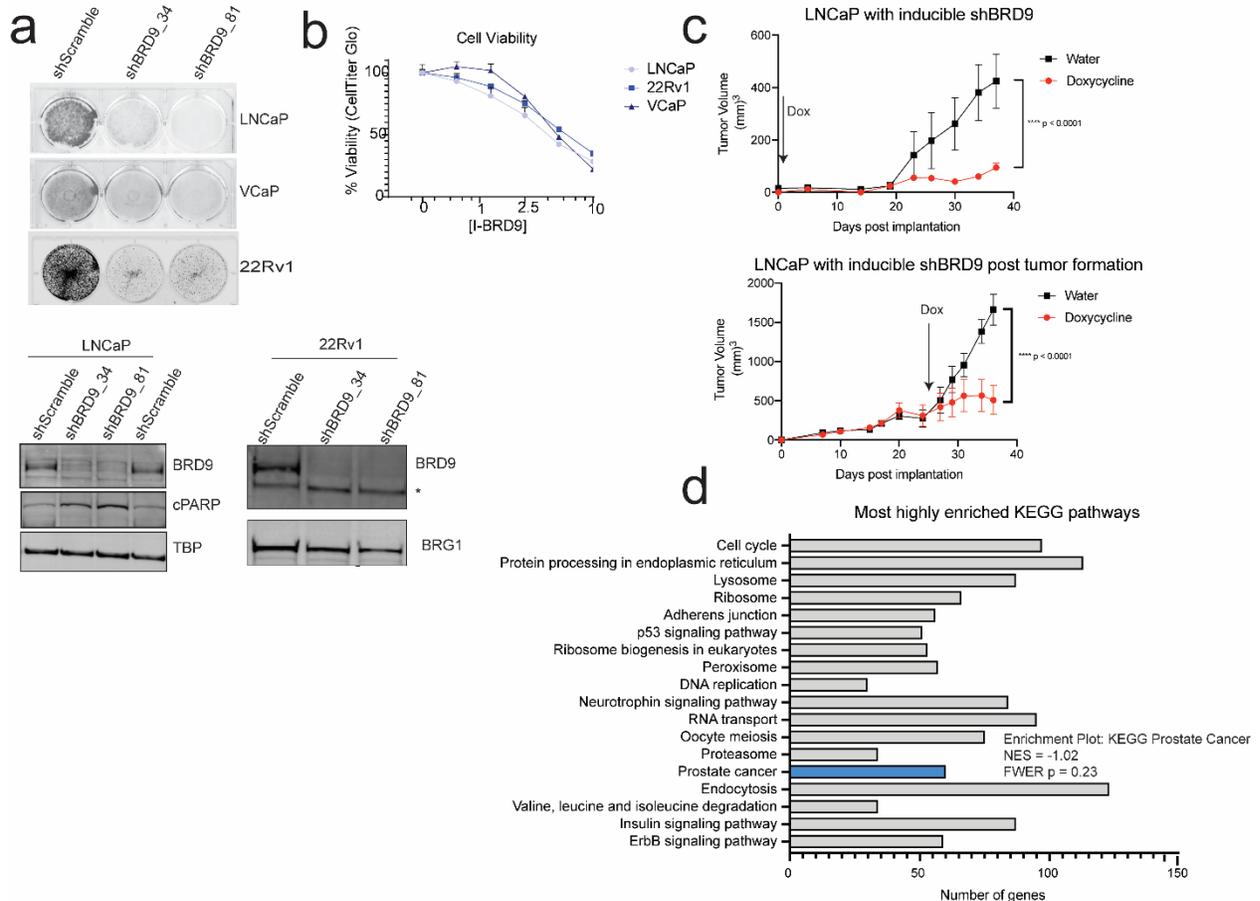
Figure 4-3 BRD9 is critical for viability of prostate cancer cells. **a** Crystal violet staining to assess the viability of prostate cancer cells. After transduction and selection, cells were plated and incubated for 12 days (LNCaP, VCaP) and 14 days (22Rv1) and stained with crystal violet. Immunoblots showing the knockdown of *BRD9*. Images are representative of two biological replicates. **b** Cell viability upon treatment with I-BRD9. Cells were plated and treated with 2-fold dilutions of I-BRD9 starting from 10 μM for 5 days. Viability is assessed by CellTiter Glo reagent. Luminescence readings are normalized to DMSO-control of each cell line and expressed as percent viability. The graph is representative of two independent experiments. Mean of three technical replicates; error bars represent standard deviation. **c** Mouse xenograft experiment with LNCaP cells. Doxycycline-treatment starts simultaneously with tumor cell injection (*top)* into nude mice or after tumor develops (*bottom*) in Nrg mice. Lines represents the average tumor size from control or knockdown groups (n = 7, in each group), error bars represent standard error of means. Statistical significance determined using multiple unpaired t tests per row (days) with the Holm-Sidak correction. ** for p<0.01; **** for p<0.0001**. d** KEGG pathway analysis of *BRD9* knockdown RNA-seq, suggesting processes such as cell cycle, DNA replication and disease such as prostate cancer were found releveant with the differentially expressed gene list.

4.3.2 *BRD9* knockdown has overlapping transcriptional effects as enzalutamide

In order to gain insight about the effect of *BRD9* knockdown, we performed RNA-seq in LNCaP cells. Based on KEGG pathway analysis, the major affected processes include on cell cycle, DNA replication and protein metabolism that can explain the viability defects observed in the cells (Figure 4-3d). Besides these generic pathways, the observation that *BRD9* knockdown has greater influence on AR-positive cell lines prompted us to speculate whether BRD9 and GBAF is involved in AR-dependent gene expression program. To test this, we compared publicly available dataset in which LNCaP cells were treated with antiandrogen enzalutamide (ENZA) for two days [116] with RNA-seq from *BRD9* knockdown LNCaP cells. We observed a correlation in gene regulation for genes differentially regulated by both shBRD9 or ENZA (Figure 4-4a) and a significant overlap in DEG lists from shBRD9 or ENZA datasets (Figure 4-4a). These genes include canonical AR target genes, such as PSA (KLK3), KLK2, EZH2, TMEFF2, SYT4, FKBP5, MYCC, CENPN.

Figure 4-4 BRD9 regulate androgen-receptor-dependent gene expression. **a** Comparison of DEGs from *BRD9* knockdown LNCaP cells and enzalutamide-treated LNCaP cells [116]. p-value for the overlap is calculated using Fisher's exact test of a contingency table generated using the number of genes expressed in LNCaP. **b** Effect of *BRD9* knockdown or inhibition with I-BRD9 on AR-target gene expression. RNA is harvested 10-day-post-transduction (including selection) or after 4 or 5-day treatment with various concentrations of I-BRD9. Data shown are representative of two independent experiment, with three technical replicates. Error bars represent standard deviation. All cells were grown in regular culture media. **c** LNCaP and VCaP cells were grown in charcoal-stripped FBS containing media (csFBS) for 2 days and pretreated with I-BRD9. Medium was switched to regular FBS containing media (FBS) with or without I-BRD9 to induce androgen-dependent gene expression. RNA was harvested after the end of indicated induction periods. Data shown are representative of two independent experiments, with three technical replicates. Error bars represent standard deviation. **d** Stable LNCaP cells having doxycycline inducible empty vector (Empty) or GLTSCR1L construct were grown in charcoal-stripped FBS containing media (csFBS) and doxycycline-treated for 2 days. Medium was switched to regular FBS containing media (FBS) with doxycycline. RNA was harvested after the end of indicated induction periods. Data shown are representative of two independent experiments, with three technical replicates. Error bars represent standard deviation. Western blot panel validating GLTSCR1L expression and its incorporation into the complex with BRD9 **e** LNCaP cells were treated with I-BRD9 for 24 h and the nuclear lysates were prepared with or without I-BRD9. BRD9 was immunoprecipitated and probed for AR and BAF155.
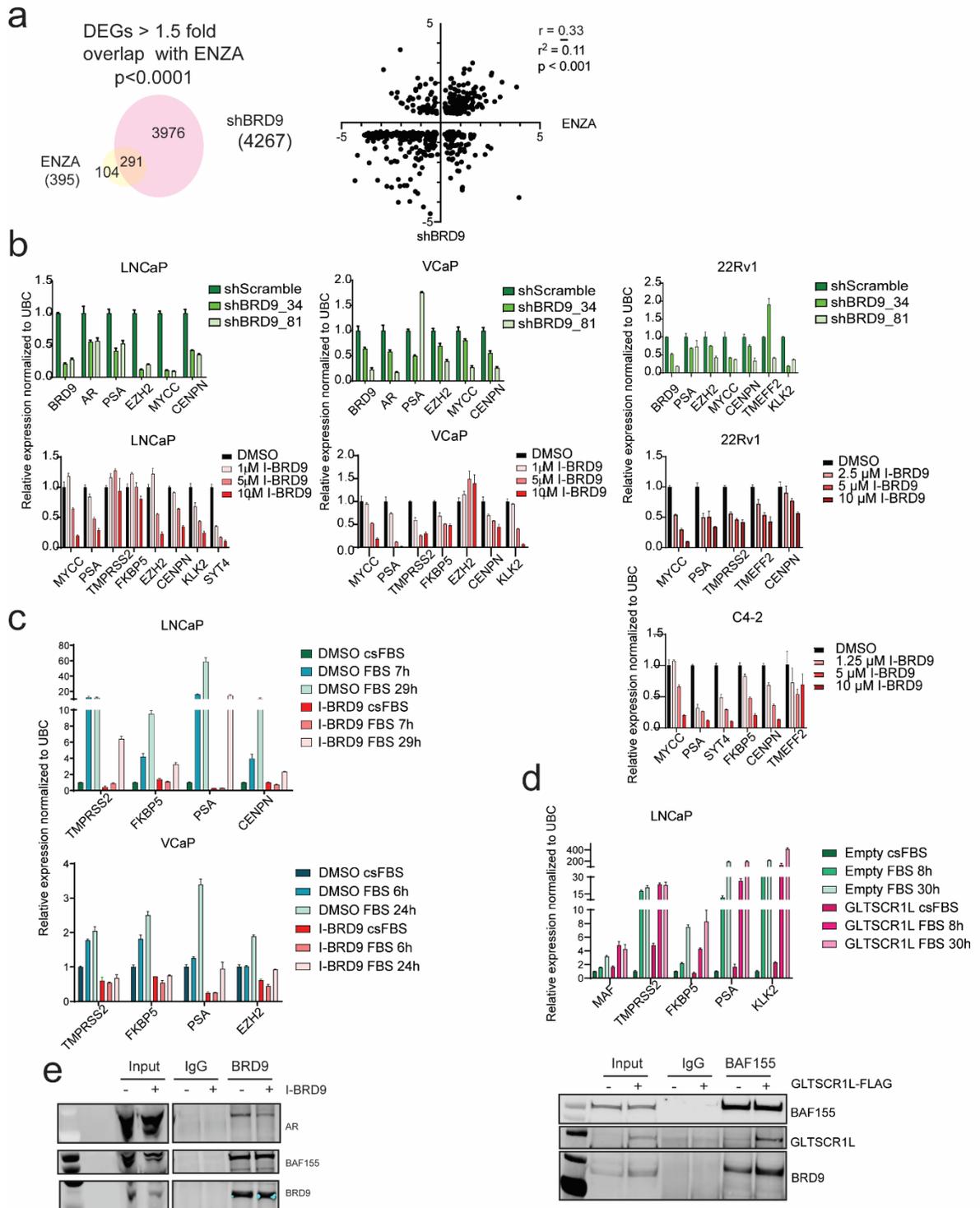
4.3.3 BRD9 can interact with AR and regulate subset of AR-target gene expression

Observing significant overlap between BRD9-regulated genes and enzalutamide-responsive genes, we tested whether generic AR target genes are affected by GBAF perturbation.

First of all, we showed that expression of a number of AR target genes and *MYC* went down upon *BRD9* knockdown in LNCaP, VCaP and 22Rv1 prostate cancer cell lines. Similarly, I-BRD9 reduced the AR target gene expression in dose-dependent manner in LNCaP, VCaP, 22Rv1 and androgen independent C4-2 cell lines (Figure 4-4b). In order to test whether BRD9 is required for dynamic regulation of androgen/AR-mediated gene expression rather than targeting the gene expression in alternative ways, we performed steroid starvation and induction experiments in the presence of I-BRD9. We first tested the utility of this experimental approach by growing the cells in complete media or steroid/androgen-deprived (charcoal stripped) FBS containing media. As expected, especially LNCaP and VCaP cells had ~50% reduced proliferation in charcoal-stripped FBS containing media (Appendix A Supplementary Figure 2a). 22Rv1 cells express both ligand responsive AR and an AR isoform lacks ligand binding domain, making the cell line partially ligand-independent [117]. That is why 22Rv1 cell count in charcoal-stripped FBS was only 25% lower than the cells growing in complete medium. Similarly, androgen starvation greatly reduced the AR target gene expression in LNCaP and VCaP (Appendix A Supplementary Figure 2b) while gene expression in 22Rv1 was only modestly affected by androgen deprivation. These observations encouraged us to implement androgen induction experiments using two-media system. In order to test the influence of BRD9 blockade, we grew the cells in charcoal-stripped FBS containing medium for 2 days with or without I-BRD9. Then, we switched the medium with complete medium to induce androgen-sensitive genes. We showed that pretreatment of LNCaP and VCaP cells with I-BRD9 delayed the induction of AR target genes upon switch from charcoal-stripped serum containing media to complete growth media (Figure 4-4c). This suggested that BRD9, thus GBAF, may directly function as a coregulator of AR and needed for both steady-state expression and rapid induction of AR target genes. Another BRD9 bromodomain probe BI-7273 and dBRD9 also caused retardation of gene induction in LNCaP cells despite having less dramatic effect compared to I-BRD9 (Appendix A Supplementary Figure 3). In addition to inhibition or depletion of BRD9, we wanted to check whether increasing GBAF abundance can cause alterations in AR-dependent gene expression. We observed that LNCaP cells with ectopic expression of GLTSCR1L can stabilize BRD9 and increase its association with core subunit BAF155, implying formation of new GBAF complexes (Figure 4-4d). We found that GLTSCR1L-expressing LNCaP cells either have higher basal level of AR target genes under androgen-deprived conditions (like TMPRSS2) or induce the expression of target genes such as MAF, KLK2, FKBP5

to a greater extent upon switch to regular media (Figure 4-4c). This implied that GBAF can boost AR activity in select target loci, as observed with some other AR coregulators [118,119]. Noticing the strong connection between AR-mediated transcription and GBAF components, we looked for possible mechanisms how GBAF is involved in AR-dependent gene expression. Since SWI/SNF complexes regulates gene expression by partnering with transcription factors, we first tested whether GBAF and AR physically associate. Immunoprecipitation of endogenous BRD9 from VCaP and LNCaP lines demonstrated that AR could associate with GBAF (Figure 4-4e). Immunoprecipitation experiments in the presence of I-BRD9 partially reduced the amount of coimmunoprecipitated AR, suggesting that bromodomain of BRD9 might contribute to this association.

4.3.4   GBAF cooperate with BET proteins

Previous evidence suggested that BET proteins and GBAF complex can physically interact in different contexts [28,29,65,120]. We and others previously showed that GBAF components GLTSCR1 and BRD9 can interact with BRD4 [28,29,65]. Extending this knowledge in prostate cancer cell lines, we showed that BRD9 can coimmunoprecipitate BRD2 and BRD4 but not BRD3 in LNCaP and VCaP cells (Figure 4-5a). We also demonstrated that the interaction between SWI/SNF complexes with BET proteins were mainly through GBAF complex (Figure 4-5a). First of all, we observed that BRD4 cannot coimmunoprecipitate BAF47, which is a dedicated subunit of canonical BAF and PBAF complexes but not GBAF. In addition, the amount of BAF155, a common subunit of all three SWI/SNF subcomplexes coimmunoprecipitated by BRD4 is significantly lower when BRD9 is depleted via shRNA or degrader, an observation that is validated by reciprocal IP of BAF155 or BRG1. Furthermore, using bromodomain inhibitors of BET proteins and BRD9, we found that the interaction between BRD9 and BET is sensitive to BET inhibitor JQ1 and BRD9 inhibitor I-BRD9 (Figure 4-5b). This observation indicated that bromodomains of both BRD9 and BET proteins –at least partially– contribute to the interaction between these proteins. Supporting this, ectopic expression of wild type BRD9 can successfully coimmunoprecipitate BRD2 and BRD4 while bromodomain-mutant BRD9 (BRD9-N216A) had significantly reduced amount of BRD2 and BRD4 coIPed (Figure 4-5b). Functionally, BET proteins have been regarded as AR-coregulator as well [68,106]. Our RNA-seq data demonstrated significant overlaps between differentially expressed genes from *BRD9* knockdown LNCaP cells

and BET inhibitor (I-BET151)-treated LNCaP cells [106], indicating that GBAF and BET proteins cooperate functionally (Figure 4-5c). These biochemical and functional foundations laid the basis for possible genomic co-targeting of BET proteins and GBAF in prostate cancer, which has been implicated in a different system [29]. In order to identify sites occupied by GBAF complex and BET proteins, we performed ChIP-seq for BRD9 and BRD4, as the conventional member of BET family proteins, in LNCaP cells. We observed 24883 BRD9 peaks and 30432 BRD4 peaks, which are mainly located at intergenic sequences. Using GSE28126 we called 10554 peaks for AR. We mapped each peak from AR, BRD4 and BRD9 to the nearest gene and looked at overlap in terms of gene identities (Figure 4-5d). We observed 3312 genes with all three factors bound in the periphery. Comparing the genes with the respective RNA-seq data, we found that of 3312 genes that are bound by AR, BRD4 and BRD9, 119 genes were differentially expressed in each of I-BET151, ENZA and shBRD9 datasets. These included the same classical AR target genes PSA, BMPR1B, and TMPRSS2. We also confirmed co-binding at the select AR target sites using ChIP-qPCR and observed a decrease in BRD9 binding with I-BRD9 (Figure 4-5e). Confirming previous genome-wide results, we did not observe significant decrease in BRD4 enrichment with I-BRD9, implying that localization of BET proteins is not dependent on GBAF, whereas GBAF localization can be dependent on BET proteins such as BRD2 and/or BRD4 (Appendix A Supplementary Figure 5c), where we did see that BET inhibitor JQ1 displaced BRD9 from the target sites.

Figure 4-5 GBAF associates with BET proteins. **a** *left* BRD9 coimmunoprecipitated BRD2 and BRD4 but not BRD3 in LNCaP and VCaP cells. *right* BRD4 principally interacts with GBAF form of SWI/SNF complexes. VCaP cells were treated with BRD9 degrader (dBRD9) for 24 hours. BRG1 and BRD4 were immunoprecipitated and eluted. BAF155 was immunoblotted as a generic subunit for all SWI/SNF complexes while BAF47 was immunoblotted as a representative subunit for non-GBAF SWI/SNF complexes (cBAF and PBAF). Similar results were obtained from *BRD9* knockdown cells (data not shown). **b** BET protein-BRD9 interaction is *partially* bromodomain-dependent. *left* VCaP cells were treated with 0.5 μM JQ1 or 10 μM I-BRD9 for 24 hours before the cells were harvested. BRD9 and BRD4 were immunoprecipitated in the presence of respective inhibitor. *right* LNCaP cells were engineered to express inducible V5-tagged wild type BRD9 (V5-BRD9) or bromodomain-mutant BRD9 (V5-BRD9 N216A). Forty-eight hours post induction with doxycycline, the cells were lysed and exogenous BRD9 was immunoprecipitated with V5 antibodies (right). **c** BRD4 and BRD9 coregulate a subset of AR targets. Correlation plot for DEGs from BET inhibitor I-BET151 treatment [105] and *BRD9* knockdown (this study) (left). Venn diagram depicting overall overlapping genes from shBRD9, I-BET151 and enzalutamide treated cells. Fold change >1.5 criterion has been implemented. p-value for the overlap (p < 0.001) is calculated using Fisher's exact test of a contingency table generated using the number of genes expressed in LNCaP. **d** Venn diagram showing the number of genes that have BRD9, BRD4 or AR peaks nearby (top). Venn diagram showing the overlap between *the genes co-bound by AR, BRD4 and BRD9* and *the differentially expressed genes common to all three datasets (I-BET151, shBRD9 and enzalutamide)*(bottom). **e** ChIP-qPCR showing the enrichment of BRD9 and BRD4 with or without I-BRD9 at defined AR target sites.

a

LNCaP | VCaP

Input IgG BRD9 + Input IgG BRD9

BRD4
BRD3
BRD2
BRD9

VCaP

Input | IgG | BRG1 | BRD4

dBRD9 − + − + − + − +

GLTSCR1
GLTSCR1L
BAF155
BAF47
BRD9
BRD2
BRD4

b

VCaP

Input | IgG | BRD9 | BRD4

JQ1   − + − − + − − + − − + −
I-BRD9 − − + − − + − − + − − +

BRD4
BRD2
BAF155
BRD9

LNCaP

Input | V5 IP

V5-BRD9   − + − − + −
V5-BRD9 N216A − − + − − +

BRD4
BRD2
V5 (BRD9)
BAF155

c

JQ1 DEG (Log2FC)

r = 0.25
r2 = 0.06
p < 0.0001

shBRD9 DEG (Log2FC)

DEGs > 1.5 fold overlap

I-BET151 (2760) 1509 1058 shBRD9 (4267)
109
84 182
20 2918

ENZA (395)

d

LNCaP ChIP-Seq
Genes with nearby binding sites

BRD4 30432 peaks

BRD9 Bound at X genes (24883 peaks)

3280 4802 2350
831 3312 399
1011

AR 10554 peaks

genes bound by all 3312

3193 119 175

co-regulated 294

e

LNCaP ChIP-qPCR

BMPR1B enhancer

0.025
0.020
0.015
0.010
0.005
% input

I-BRD9 − + − + − +
IgG   BRD9   BRD4

KLK3 enhancer

0.05
0.04
0.03
0.02
0.01
% input

I-BRD9 − + − + − +
IgG   BRD9   BRD4

4.3.5 The redundant subunits GLTSCR1 and GLTSCR1L have modest effects on the prostate cancer line, LNCaP

Our initial report prompted a potential role for GLTSCR1 and GLTSCR1L in prostate cancer cell line PC3, whose proliferation was reduced ~40% upon knockouts. However, for LNCaP cells, we did not observe the similar trend such that except from a transient retardation in the growth (Figure 4-2a), the growth defect sourced from *BRD9* knockdown was not comparable to the knockout of either of the paralogous subunits. RNA-seq from *GLTSCR1* and GLTSCR1L demonstrated fewer differentially expressed genes and smaller magnitude of differential expression compared to *BRD9* knockdown data (Figure 4-6a), which is comparable with the strength of the growth phenotypes observed among the three manipulations.

Based on the RNA-seq, we realized a subset of AR target genes were among the DEGs of *GLTSCR1* and *GLTSCR1L* KO datasets (Figure 4-6b). Still, the fold changes associated with these DEGs were smaller than they were in *BRD9* knockdown dataset. In terms of genomic localization, we performed ChIP-seq with BRD9 antibody in control, *GLTSCR1* KO and *GLTSCR1L* KO cells (Figure 4-6b). We found that most of the BRD9 sites were common in all three cell lines, while we observed a number of BRD9 target sites that were unique to either conditions, meaning BRD9 sites were specifically lost or gained upon knocking out *GLTSCR1* or *GLTSCR1L,* which is in agreement with the co-IP experiments that demonstrated BRD9 is still GBAF– and BRD4– associated in *GLTSCR1* or *GLTSCR1L* knockout cells. Overall, compatible with the expression data, we did not observe a big difference in the total number of BRD9 peaks across the cell lines once individual paralogs were knocked out.
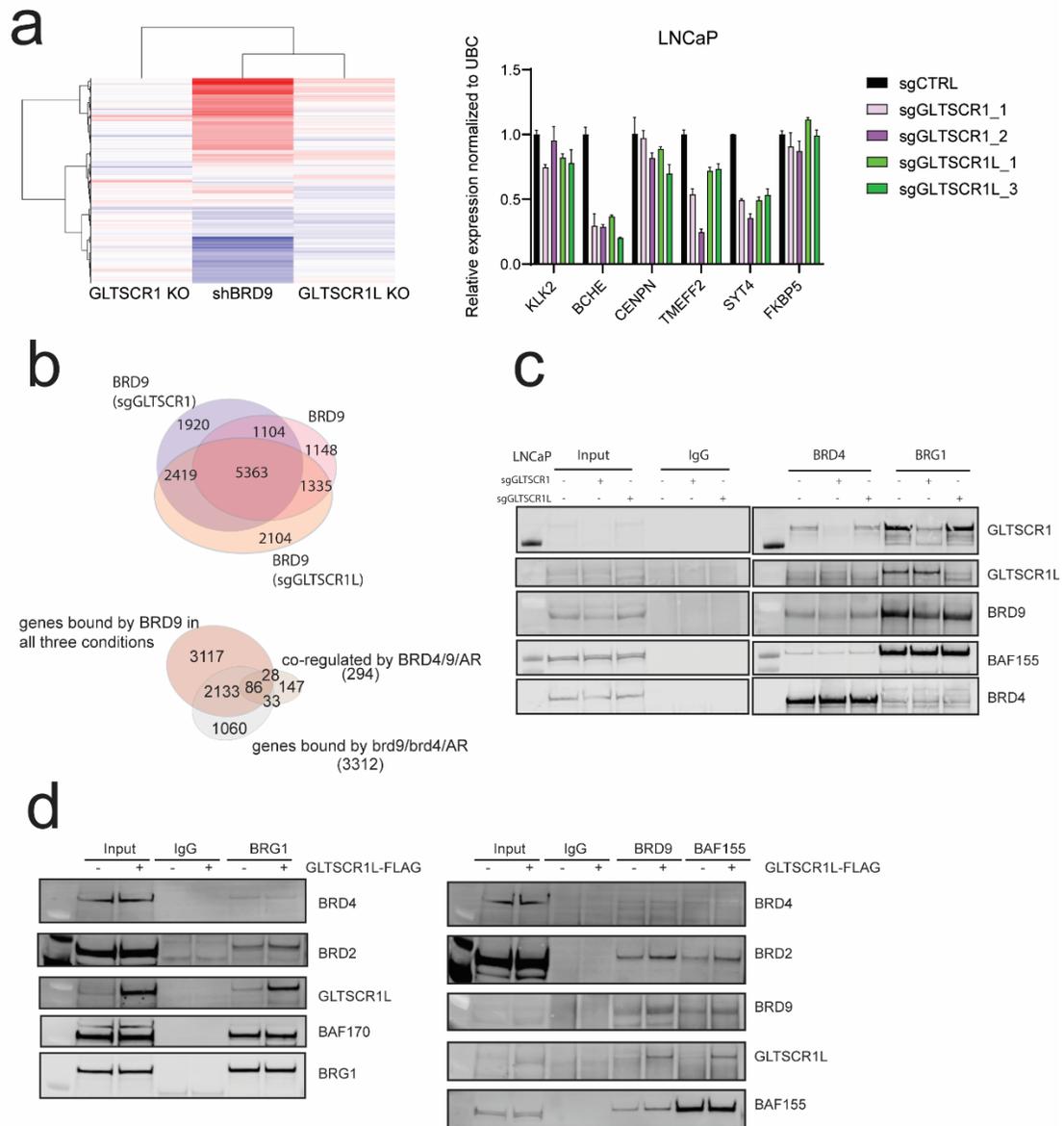
Figure 4-6 GLTSCR1 and GLTSCR1L have mild effect on prostate cancer cells. **a** Heat map depicting the clustering of DEGs from *BRD9* knockdown, *GLTSCR1 KO* and *GLTSCR1L KO* cells. Relative expression of select AR target genes in paralog knockout LNCaP cells was assessed via RT-qPCR. Data shown are representative of two independent experiments with technical replicates n=3. Error bars show standard deviation. **b** Venn diagram depicting the overlaps among BRD9 ChIP-seq peaks in control, *GLTSCR1* KO and *GLTSCR1L* KO cells. **c** *GLTSCR1* or *GLTSCR1L* knockout did not significantly compromise BRG1–BRD9 or BRD4– BRD9 interaction. **d** Exogenous GLTSCR1L expression enhances association of BRD2 but not BRD4 with BRG1 or BAF155. LNCaP cells were engineered to express inducible FLAG-tagged wild type GLTSCR1L were treated with doxycycline for 2 days to express GLTSCR1L. BRG1 (left) or BRD9 and BAF155 (right) were immunoprecipitated and the eluted proteins were run on gel.

The superiority of BRD9-related phenotypes and the refractoriness of BRD9 localization in paralog-knockout setting promoted us to revisit the complex dynamicity in prostate cancer cell lines. Although BRD9 and GLTSCR1/1L paralogs are acting in the same complex, the observation that GLTSCR1 and GLTSCR1L can –at least partially– substitute the loss of one another may render the complex still intact in the absence of either paralog. In agreement with VCaP data (Figure 4-5a), we observed that in LNCaP cells GLTSCR1 is the main paralog that BRD4 interacts in the context of GLTSCR1, considering its enrichment over the input (Figure 4-6c). We failed to generate cell lines which are completely deficient in both subunits and observed partial expression from either paralog in the resulting co-transduced cell lines (data not shown), which implicated possible depletion of the dual knockout cells from the culture while recovering from transduction. This might imply a dependence to the overall "dosage" of these paralogous subunits. In terms of BRD9 dependence, previous data showed that GLTSCR1–BRG1 interaction (referred as "complex incorporation") is reduced by BRD9 degrader (dBRD9) (Figure 4-5a) or *BRD9* knockdown (data not shown) while that of GLTSCR1L is even slightly enhanced. Interestingly, the same contrasting observation was made in a recent SWI/SNF knockout screen [121]. Taking together, these observations reflect the heterogeneity of GBAF complex assembly depending on the paralog incorporation.

In terms of GBAF-BET interaction, we found that VCaP and LNCaP cells both BRD2 and BRD4 interact more efficiently with GLTSCR1 than GLTSCR1L in endogenous levels (Appendix A Supplementary Figure 4a). Despite poor enrichment at endogenous levels, once expressed exogenously, GLTSCR1L can interact with BRD2 and BRD4 in HEK293T cells and LNCaP cells (Appendix A Supplementary Figure 4). With GLTSCR1L expression, we observed that BRG1 and BAF155 associates more with BRD2 and slightly less with BRD4, suggesting that GLTSCR1L-GBAF complex has slight preference to BRD2 (Figure 4-6d). It appears that relative abundancies of the paralogous subunits of GBAF can be a determinant of the type of BET protein that GBAF interacts. The functional consequences of this slight BET protein preference need further investigation. Taking the previous data into account as well, it appears that BRD9 is the most critical GBAF–specific subunit for BET protein interaction.

4.4 Discussion

In this study, we report that recently discovered, novel type of SWI/SNF subcomplex (GBAF or ncBAF) is critical for maintenance of prostate cancer cells by partnering with AR and BET proteins. We dound that knocking down the dedicated GBAF subunit *BRD9* led to growth inhibition in various prostate cancer cell lines. Similarly, small molecule bromodomain inhibitor I-BRD9 slowed down the growth of these cell lines. Immunodeficient mouse xenografts of LNCaP gave rise to smaller tumor sizes upon *BRD9*-depletion, confirming the *in vivo* validity of BRD9 dependence. We observed that BRD9 can interact with AR and inhibition and knockdown of BRD9 led to reduction in AR-target gene expression. We also reported a functional link between GBAF and another group of AR coregulators, bromodomain and extra-terminal domain proteins (BET). We demonstrated significant overlap between transcriptional effects of *BRD9* knockdown and BET inhibitor treatment, implying GBAF and BET proteins cooperate. We reported that BET proteins BRD2 and BRD4 can interact with mainly GBAF form of BAF complexes and that BRD9- BET protein interactions are partially bromodomain dependent. In addition, we found that BRD4 and BRD9, together with AR have overlapping genomic localizations. Taken together, our study revealed GBAF complex as another coregulator of AR that works together with BET proteins.

We and others previously characterized this new complex in different cell types including synovial sarcoma, malignant rhabdoid tumors, myeloid leukemia and embryonic stem cells [27,28,30,31,33], demonstrating that the complex incorporates specific subunits of GLTSCR1, GLTSCR1L (GLTSCR1L) and BRD9 but not well-established core SWI/SNF subunits such as BAF57, BAF47, ARID1/2 or BAF170. As a recently established complex, the little was known specifically about the roles of GBAF in SWI/SNF-intact settings. Despite being one of the highly mutated factors in cancer [19], SWI/SNF mutations are rare in prostate cancers, which may imply dependence of prostate cancer cells to intact SWI/SNF complexes. Previous studies have implicated the critical roles of SWI/SNF chromatin remodelers in prostate cancer especially the ones that bears high-incidence mutations such as *PTEN* deletion and *ERG* fusions. For example, *PTEN*-deleted prostate cancers have more dependence to BRG1 [113]. In addition, ETS transcription factor ERG is dependent on SWI/SNF complexes for proper chromatin targeting and gene activation [122].

Androgen signaling is the major vulnerability of prostate cancers. Being the major organ responsive to systemic androgens, it serves as an attractive target for prostate cancer treatment. Although several new-generation anti-androgens can target AR more efficiently, acquired resistance to androgen deprivation therapy persists as the major problem. Previous reports identified AR coregulators that AR signaling is dependent on for proper target gene expression and several SWI/SNF subunits were also implicated as AR coregulators [104,110,111]. Our current report expanded the roles of the SWI/SNF members in AR-dependent transcription, by demonstrating that select AR target genes are downregulated upon inhibition of BRD9 bromodomain and -at varying levels- upon depletion of BRD9, GLTSCR1 or GLTSCR1L.

SWI/SNF core subunits such as BRG1, BAF155 were reported to interact with BET proteins BRD2 [123,124]; BRD3 [125]; BRD4 [77] in different systems. However, the interactions of individual subcomplexes with BET proteins have not been elucidated in detail. Previously, GLTSCR1 has been assigned as BRD4 ET-domain interacting protein [65]. BRD9 has also been shown to engage BRD4 ET-domain via its N-terminal region [120] that do not cover neither bromodomain or C-terminal domain required for GBAF complex incorporation [27]; while another report suggested the requirement of intact BRD9 bromodomain for BRD9-BRD4 association using inhibitors [29]. We have observed in endogenous IPs of BRD9 and BRD4 that either JQ1 or I-BRD9 treatment can decrease the associations between BRD9-BRD4 or BRD9-BRD2. Additionally, bromodomain-mutant BRD9 (BRD9-N216A) that cannot engage acetyl-lysine moiety [33] could not efficiently interact with BET proteins, further suggesting that BRD9 bromodomain recognizes BET proteins. Further research is needed to identify acetylation events on BET proteins that are critical for GBAF interaction.

GLTSCR1 and GLTSCR1L are two dedicated, mutually-exclusive subunits of GBAF complex. Our RNA-seq analysis demonstrated that *GLTSCR1L* knockout has more differentially expressed genes than *GLTSCR1* knockout. In addition, DEGs from *GLTSCR1L* knockout cells have better overlap with *BRD9* knockdown cells compared to those from *GLTSCR1* and *BRD9* datasets. Using co-IP assays, we tried to assess the intra-complex dependencies of GBAF-specific subunits. We observed that *BRD9* knockdown led to partial reduction in protein level and BRG1 or BAF155 association of *GLTSCR1*; whereas *GLTSCR1L* has enhanced association with *BRG1* when BRD9 is depleted. In addition, knocking out neither *GLTSCR1* nor *GLTSCR1L* affected BRD9 levels or its association with BRG1 or BAF155. Likewise, in terms of GBAF-BET

interactions, the subunits have differential roles. We observed that BRD9 is critical for BRD2/4 interactions. Although GLTSCR1 and GLTSCR1L can be co-IPed by BRD2 and BRD4, the former is more enriched. In addition, despite increased levels of protein in *BRD9* depleted cells, interaction between GLTSCR1L and BRD4 decreases -even when GLTSCR1L is exogenously expressed, confirming that BRD9 is critical for GBAF-BET interaction. Considering also the drastic differences in the growth phenotypes, we think that BRD9 is the most critical GBAF-specific subunit in the complex for prostate cancer. Based on the *intra-complex* dependencies, further investigation is needed for the functionality of GBAF without GLTSCR1/1L paralogs and GLTSCR1L-GBAF without BRD9. We think that some of the discrepancies between transcriptional effects upon individual subunit depletions can be explained by the acquisition of these *partial* complexes.

AR target genes are often controlled by AR binding to distal regulatory elements [126–128]. The mechanism of gene regulation by AR at distant regulatory sites is hypothesized to be through looping. However, the precise role of regulators of global chromatin conformations, such as CTCF, has not been understood for long-range AR-dependent transcriptional regulation. One study suggested that a subset of AR-responsive genes that are located within two successive CTCF sites (referred as CTCF blocks), which are enriched for AR, FOXA1 and H3K4me2, tend to be more highly expressed and be more relevant with cancer-related pathways [129]. CTCF sequestration is critical for efficient long-range communication for the other transcriptional events as well. Previous research reported that BRD2 and BRD3 but not BRD4 colocalizes with CTCF [130,131]. It is suggested that BRD2 localization is dependent on CTCF binding to its cognate motif, where BRD2 binding provides proper insulation at the boundaries such that long-range interactions do not "leak" across the boundaries, which otherwise causes aberrant contacts and aberrant gene regulation. Comparing previous BRD2 and BRD4 ChIP-seq from VCaP cell line [68] with CTCF ChIP-seq (ENCODE), we observed that BRD2 shares more sites with CTCF than BRD4 does, although BRD2 and BRD4 enriched sites are highly overlapping (Appendix A Supplementary Figure 5a) contrast to previous studies in Th17 [131] and erythroblast cells [130]. Although a previous study investigated BET protein roles at the level of AR localization and AR-dependent gene expression, potential CTCF-mediated role of BRD2 on AR-regulated gene expression is yet to be elucidated. Using ChIP-qPCR, we demonstrated that BRD2 is enriched at CTCF sites nearby the AR-target genes, such as FKBP5, and TMPRSS2 (Appendix A Supplementary Figure 5c). In

addition to possible BRD2 function at CTCF sites for AR-dependent gene expression, it is unclear whether GBAF is also involved in this type of indirect regulation. Previous research suggested the specific GBAF occupancy at CTCF motifs, while only a small fraction of BAF or PBAF peaks is enriched at CTCF sites [27,29]. We detected BRD9 enrichment at the very same CTCF sites of BRD2 enrichment. Mechanistically, we observed that BRD9 localization is impaired by both JQ1 and I-BRD9, while BRD2 enrichment only decreased by JQ1, suggesting that GBAF localization is dependent on BRD2 binding at CTCF binding sites as well. Furthermore, we identified a putative topologically associating domain (TAD) from LNCaP Hi-C dataset (ENCODE) bordered with BRD9-CTCF colocalization site. Interestingly, this domain encompasses a number of genes that are significantly differentially expressed in our *BRD9* knockdown RNA-seq dataset (data not shown). In short, co-binding of BRD2 and GBAF might regulate gene expression through defining the chromatin domain ends and chromatin looping. In agreement with this hypothesis, we detected that bromodomain-mutant BRD9 fails to pulldown BRD2 as well as CTCF while wild type BRD9 coimmunoprecipitates both (Appendix A Supplementary Figure 5b). This suggested that BRD9 interaction with CTCF might be direct or mediated by BET protein BRD2 (or possibly other isoforms in different sites) but dependent on its bromodomain. Further research is needed to elucidate the global effect and function of GBAF enrichment at the CTCF sites. Major questions remain as whether GBAF has similar chromatin remodeling activity as other BAF subcomplexes and whether GBAF-mediated gene regulation is mediated by chromatin accessibility as suggested for other BAF complexes [26]. Based on CTCF colocalization suggested by other studies, GBAF may act through distinct mechanisms such as CTCF-dependent enhancer-promoter communications or establishment of topologically associating domain (TAD) boundaries.

4.5   Contributors

Mouse injections, tumor size measurements and tumor collection were performed in Biological Evaluation Core by Sandra Torregrosa-Allen and Melanie Currie. ChIP-seq and RNA-seq data was analyzed by Dr. Sagar Uttukar, Dr. Emily Dykhuizen and Dr. Benjamin Carter. All *in vitro* experiments were performed by Aktan Alpsoy.

CHAPTER 5.   FUTURE DIRECTIONS

5.1   RNA-dependent targeting and functions of Chromobox homolog (CBX) proteins

5.1.1   Introduction

Polycomb group proteins are involved in spatial and temporal regulation of gene expression at chromatin level. These proteins are crucial components of distinct transcriptional programs that drive differentiation, self-renewal and establishment of cellular identity. Multiple line of evidence suggested their dysregulation as a cause of several disease states including cancer. In higher eukaryotes, polycomb groups form two main multimeric complexes, namely polycomb repressive complex 1 (PRC1) and polycomb repressor complex 2 (PRC2). The activities of these complexes create transcriptionally non-permissive environment through related but biochemically distinct mechanisms such as deposition of repressive histone marks, chromatin compaction etc. Basically, the two complexes act in concert to provide gene repression: PRC2 is targeted to specific genomic locations through sequence-specific DNA-binding proteins such as transcription factors or *cis*- or *trans*-acting RNA molecules including long-noncoding RNAs. EZH2, the catalytic subunit of PRC2, deposits trimethylation mark on lysine 27 of histone H3 (H3K27me3). This mark acts as a docking site for PRC1: Methyl-lysine reader chromobox homolog (CBX) subunit of PRC1 recruits the complex to H3K27me3 site and through E3 ligase activity of RING1 subunit, PRC1 complex monoubiquitinates lysine 119 of histone H2A (H2AK119ub1) (Figure 5-1). PRC1 is unique in the sense that unlike PRC2, the complex is highly heterogeneous because of the presence of multiple paralogous subunits. In addition to obvious heterogeneity in canonical PRC1, mammalian systems also possess variant PRC1 complexes, which are marked by the absence of CBX subunits. Interestingly, these variant complexes are the principle form of PRC1 in certain contexts [132,133]. This high heterogeneity is the main reason behind the context-dependent roles of PRC1 complexes, which can even be opposing. Current questions include (1) how different PRC1 complexes share out their biological roles such as gene repression, short-range chromatin compaction or higher-order chromatin packing; (2) which paralogous subunits are inevitable for each of these biological functions; (3) what confers the site-specific localization and targeting of CBX-PRC1 complexes.
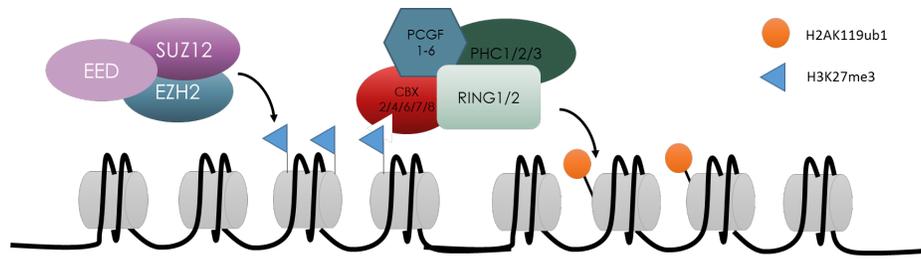
Figure 5-1 Basic architecture of canonical PRC1 and PRC2 complexes and the canonical pathway of polycomb recruitment. Combination of paralogous subunits creates multiple possible PRC1 complexes, which have unique and redundant functions. In terms of recruitment, PRC2 complex (EED, SUZ12, EZH2) deposits H3K27me3 marks, which act as docking site for canonical PRC1. CBX paralogs of PRC1 have chromodomains as H3K27me3 readers. Through its E3 ligase catalytic subunit RING1A/B, PRC1 deposits H2AK119ub1 mark.

Recent studies shed light into specific functions that can be attributed to canonical versus variant PRC1 complexes. For instance, H2AK119 ubiquitination is mainly catalyzed by the RING1 in the context of variant PRC1 [132]; while polycomb-repressed domains are mainly decorated with canonical PRC1, which compacts the chromatin in short-distance independent from ubiquitination activity [134]. This demonstrated that even if they constitute a minor fraction of total PRC1 in certain contexts, canonical PRC1 is required to compact and silence the chromatin. However, how canonical PRC1 localizes to its target sites is not completely understood. Classical model of PRC1 recruitment involves CBX proteins. Although the cognate ligand for CBX proteins is trimethyllysine moiety on H3K27, *in vitro* studies suggested that this binary interaction is a low-affinity interaction for multiple CBX proteins. Besides, some CBX proteins such as CBX7 have even more affinity to H3K9me3 [135]. One would expect that if chromodomain-H3 tail interaction was the only force governing CBX-PRC1 binding to chromatin, then, CBX proteins should have been distributed across the genome based on their affinities toward the trimethyllysine marks when expression levels are similar. Indeed, contrary to its higher affinity for H3K9me3, the majority of CBX7-enriched sites overlaps with H3K27me3 domains [136]. Similar reports also suggest colocalization of CBX6 and CBX8 with H3K27me3 even when CBX7 expression is higher than either [137]. Additionally, it is demonstrated that chromodomain-truncated CBX proteins can still interact with chromatin [138]. All these data suggested that other factors in addition to chromodomain-mediated methyl-lysine recognition are required for the proper interaction with chromatin. One promising factor is reported as non-coding RNAs. In early 2000s, it was discovered that chromodomain modules can interact with RNA [139]. Particularly, chromodomains

of mouse polycomb CBX proteins can interact with RNA [140]. In addition to these *in vitro* studies, a non-coding RNA associated with *INK/ARF* locus, regarded as *ANRIL* was found to be a direct interaction partner of CBX7 [141], an interaction that fine-tunes histone binding of CBX7 chromodomain and tethers CBX7 to the locus and potentiates the repression. Although it was reported for CBX7-PCR1 targeting and repression, we lack the information whether RNA-mediated targeting and histone affinity fine-tuning are generic for other CBX proteins and whether it is a direct and active recruitment system. Based on the current reports we hypothesized that RNA is a crucial component of PRC1 targeting at multiple target sites and PRC1-mediated gene repression.

5.1.2 Preliminary results

CBX paralogs (CBX2, CBX4, CBX6, CBX7, CBX8) are methyllysine reader subunits of polycomb repressor complex 1. Assumably, they are regarded as H3K27me3 readers, due to colocalization of PRC1 and PRC2 as well as CBX proteins and H3K27me3. However, multiple *in vitro* reports showed that the chromodomains of certain CBX paralogs have higher affinity for H3K9me3 compared to H3K27me3. Indeed, some chromodomains such as those of CBX6 and CBX8 are weak-interactors of H3K27me3. In addition to being histone-methyllysine readers, chromodomains turned out to be RNA interactors. Mouse CBX chromodomains except CBX2 can bind RNA *in vitro* [140]. Another report suggested CBX7 interacts with ANRIL, an lncRNA associated with *INK/ARF* locus. Interaction with ANRIL directs CBX7-containing PRC1 complex to the locus [141]. In order to extend our knowledge for the RNA-CBX interactions and their functional roles, we studied two probable models: We think that RNA may act as targeting factor for CBX proteins, which have moderate to low affinity to their cognate ligand H3K27me3. Secondly, RNA may mediate protein-protein interactions that can target CBX-loaded PRC1 complexes to distinct genomic sites and coordinate the gene repression.
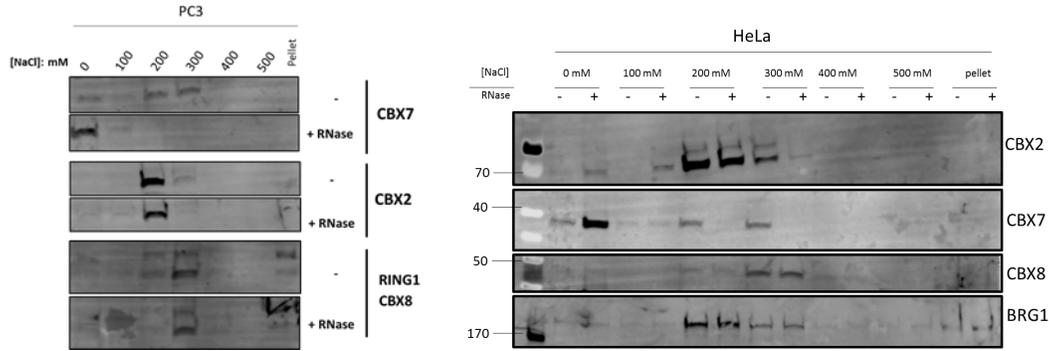
Figure 5-2 CBX7's chromatin binding is RNase-sensitive. Serial salt extraction assay suggested that CBX7's interaction with bulk chromatin is loosen with RNase-treatment. Other paralogs CBX2, CBX8 as well as CBX4 and CBX6 (data not shown) were not affected by RNase treatment. The blot is representative of at least 3 biological replicates and 3 other cell types (including HEK293T, T98G, and mESC E14Tg2a).

Firstly, we tested if interaction of CBX proteins with bulk chromatin is affected by loss of nuclear RNA. We treated the nuclei prepared from HEK293T, PC3 prostate cancer and HeLa cervical cancer cell lines with RNase A/T1 and then serially extracted the nucleoplasmic or chromatin-bound proteins using an extraction buffer with increasing concentration of salt. We observed that RNase treatment mobilized the major fraction of chromatin-bound CBX7 but not the other CBX paralogs (Figure 5-2).

Second, we tested whether active transcription is critical for engagement of CBX7 on chromatin. For that, actinomycin D is used to block transcription for 8h and the cells were fractionated. It turned out that chromatin-bound CBX7 decreased upon actinomycin treatment, providing another evidence for RNA-dependence of CBX7 for chromatin engagement (Figure 5-3).
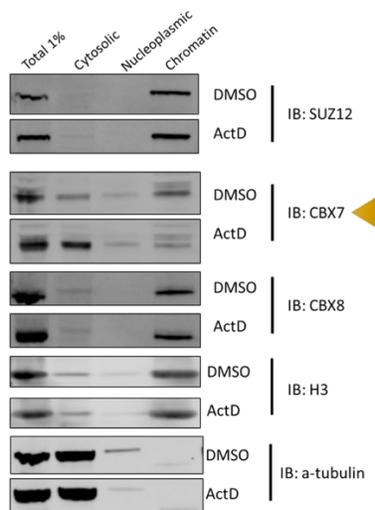
Figure 5-3 Actinomycin D-mediated blockade of transcription dislodges CBX7 off the chromatin. HEK293T cells were treated with actinomycin D for 8h before the cells were lysed for subcellular fractionation.

ANRIL is one of the best characterized non-coding RNA partner of CBX7 that recruits PRC2 and PRC1 to INK/ARF locus to repress cell cycle arrest protein p16, p15 [141]. To test if ANRIL is the sole reason behind RNA-dependence of CBX7 localization, we knocked it down in PC3 cells. *ANRIL* knockdown immediately resulted in growth arrest, as expected, and p15 was upregulated. Serial salt extraction with or without RNase in control cells and ANRIL knockdown cells (shANRIL) demonstrated that knocking down ANRIL did not mobilize chromatin-bound CBX7 (Figure 5-4), suggesting that binding to and being targeted by ANRIL are not enough to explain global RNA-dependent chromatin binding of CBX7.
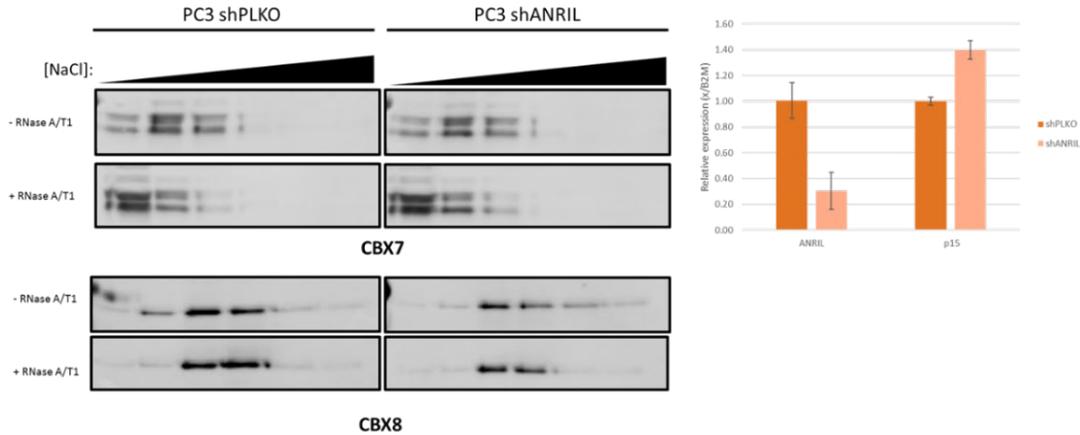
Figure 5-4 *ANRIL* is not the only RNA species responsible for global RNA-dependence of CBX7-chromatin interaction. Serial salt extraction assay showed that CBX7 elution profile with RNase is comparable between control cells and *ANRIL* knockdown cells (left). *ANRIL* knockdown is validated through RT-qPCR. As expected p15 tumor suppressor increased upon *ANRIL* knockdown.

Next, we tested if there is direct or indirect interaction between RNA and CBX paralogs by utilizing UV-crosslinking and immunoprecipitation approach (CLIP). Briefly, we UV-crosslinked the PC3 cells, lysed the nuclei and immunoprecipitated CBX7, CBX8 and hnRNPC1/C2 (as a positive control). Following high stringency washes, we labeled the associated RNA in the immunoprecipitates by radioactive ATP in a T4 polynucleotide kinase reaction. The CLIP experiment suggested that in PC3 cells CBX7 lane has slight labeling at the spot coinciding with the monomeric molecular weight (Figure 5-5, yellow arrowhead). Consistent with HEK293T CLIP (data not shown), another discrete and more robust signal was observed at around 70 kDa. This may be reflective of an RNA-binding protein that also interacts stably with CBX7. Going through Biogrid database, we observed couple of RNA binding proteins that were experimentally shown to interact with CBX7. Expanding the list from the literature we had candidates such as MOV10, DDX5, U2AF2 that are known to be nuclear. Although our CLIP failed to detect signal from MOV10 due to failure in pulldown, it is unlikely that the signal at 70 kDa is coming from MOV10 due to its higher molecular weight while DDX5 and U2AF2 have similar masses around 65-70 kDa.
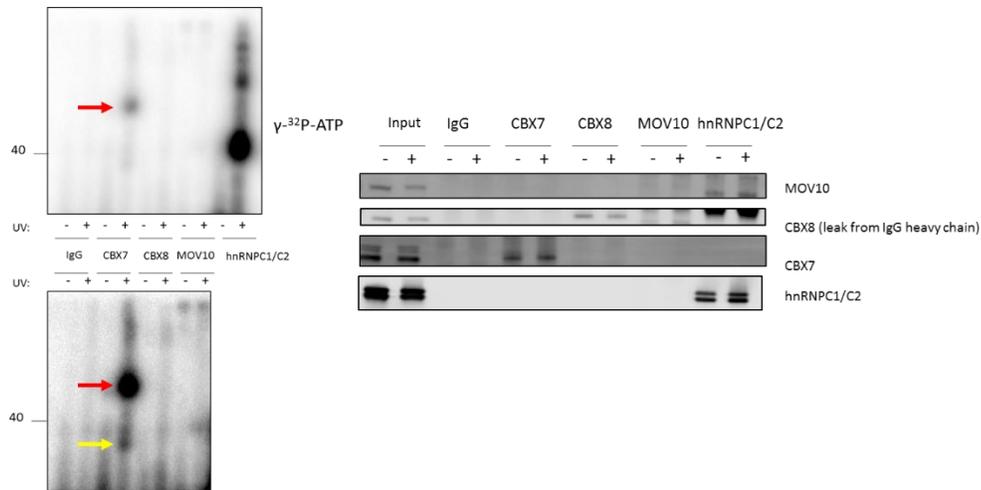
Figure 5-5 CLIP assay showed that CBX7 can directly and indirectly interact with RNA. Short exposure (top left) and high exposure (lower left) on phosphoimager demonstrated signal in CBX7 lane. We did not detect signal from CBX8 lanes. Signal at the native monomeric CBX7 size is marked with yellow arrowhead. Labeling at ~70 kDa possibly reflects a protein partner of CBX7 that is capable of interacting with RNA. hnRNPC1/C2 was used as positive control and the signal blob at 40 kDa belongs to native, monomeric protein. Note that MOV10, as CBX7 interacting RNA binding protein [142] was also checked but we failed to pull down MOV10 (IP blot, right panel), leaving MOV10 data inconclusive.

We started our testing by dead-box helicase 5 (DDX5), an RNA helicase which is involved in RNA metabolism, transport and stability as well as transcriptional regulation [143]. DDX5 was also ChIPed in Th17 cells [144], suggesting it can interact with chromatin, providing another line of support for its involvement in gene regulation at chromatin level. DDX5 was shown to interact with RING1B and CBX7 in mouse embryonic stem cells [145] and with PRC2 in hepatocellular carcinoma cells [146]. As a proof-of-concept, we seek whether CBX7 and DDX5 interact in other contexts as well: In HEK293T cells, DDX5 was detected in CBX7 IP and exogenously expressed CBX7 was detected in DDX5 IP, providing evidence that DDX5 and CBX7 interact (Figure 5-6).
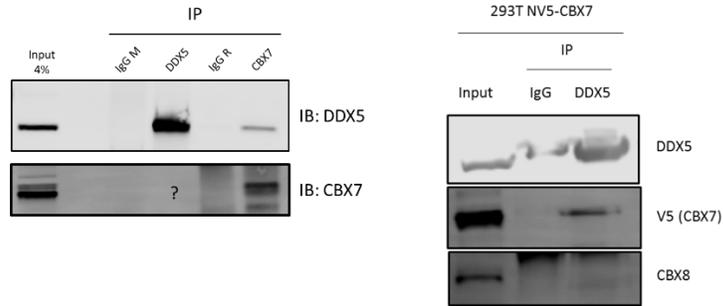
Figure 5-6 DDX5 and CBX7 interact in HEK293T cells. Endogenous IP of CBX7 pulls down detectable amount of DDX5, while the reciprocal IP fails CBX7 detection, possibly due to differences in the abundance of the proteins (left). DDX5 pulls down exogenously expressed V5-tagged CBX7 from HEK293T cells (right).

After validation of interaction, we tested whether RNA-dependent chromatin localization of CBX7 is dependent on DDX5. We generated CRISPR/Cas9 knockout HEK293T cells with three different guide RNA sequences and screened around 10 colonies from each set and selected one clone for downstream analysis. We performed serial salt extractions in control and *DDX5* knockout cells with or without RNA. We observed that neither RNA-dependence nor elution profile was affected by *DDX5* knockout, suggesting that DDX5 was not involved in regulation of CBX7's chromatin binding (Figure 5-7). Still, the essence of DDX5-CBX7 interaction and the functional consequences of this partnership remain to be elucidated.
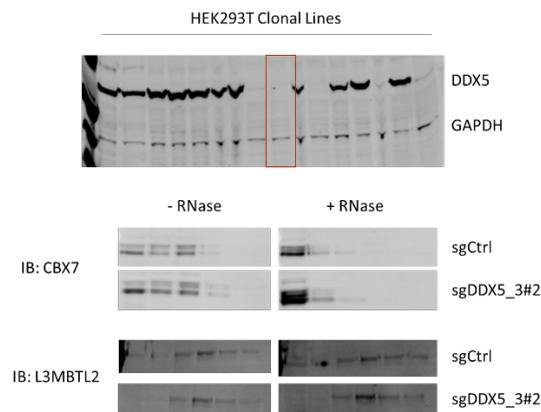
Figure 5-7 DDX5 is not involved in RNA-mediated chromatin interaction of CBX7. *DDX5* KO clones were screened for DDX5 expression; for the downstream assays the clone in red box was picked (sgDDX5_3#2). Another clonal line was picked from the control guide treated cells (sgCtrl) (upper panel). Loss of DDX5 did not affect chromatin binding of CBX7. non-canonical PRC1 subunit L3MBTL2 was used as control (lower panel).

95

We then turned our attention to another candidate factor, U2AF2. Besides DDX5, there are spliceosome associated proteins that were reported to coimmunoprecipitate with CBX7. In early studies alternative splicing has been associated with the histone marks, including H3K27me3 [147,148], which "flags" exon inclusion vs. skipping. The factor we moved forward with was U2AF65 (U2AF2, U2 Small Nuclear RNA Auxiliary Factor 2), which was shown to interact with CBX7 [149]. We first verified that CBX7 interacts with U2AF2 together with a distinct spliceosome component, SF3B2, in an RNA-independent manner (Figure 5-8). To understand which region of CBX7 is responsible for interaction with spliceosome we used domain-swapping approach where we swapped the chromodomains between CBX7 and CBX8. Transient expression followed by coimmunoprecipitation showed that chromodomain is not the main contributor of this interaction. Supporting the chromodomain-independent nature of the interaction, chromodomain-specific peptide-based probe [150] developed by Prof. Frye laboratory at University of North Caroline, Chapel Hill did not affect the interaction between U2AF2 and CBX7 (Figure 5-8).
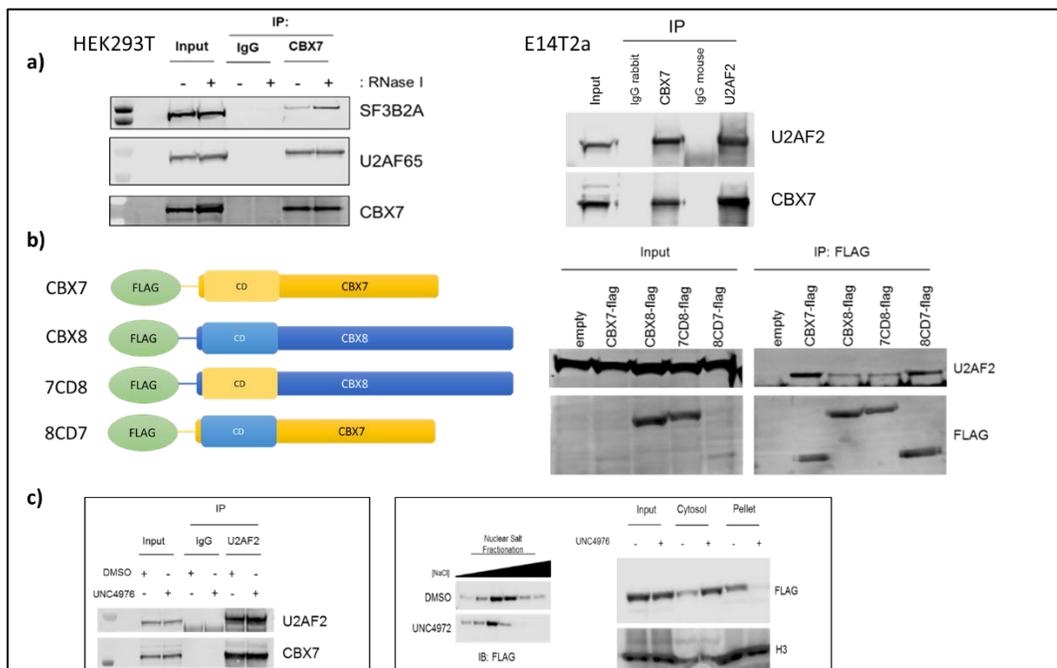
Figure 5-8 U2AF2 (U2AF65) interacts with CBX7. **a** CBX7 pulls down spliceosome component SF3B2A and auxiliary factor U2AF2 in RNA-independent manner in HEK293T cells. U2AF2 and CBX7 interaction was confirmed in another cell line through reciprocal IPs. **b** Schematic illustration of chromodomain swapping between CBX7 and CBX8. CBX7 extra-chromodomain region appears to be responsible for U2AF2 interaction. **c** UNC4976, a CBX-specific peptide-based probe UNC4976 [150] did not affect CBX7-U2AF2 interaction (left), despite mobilizing exogenously expressed CBX7 off the chromatin pellet (right).

As a follow-up experiment, we are planning to perform IPs with truncated proteins with loss of chromodomain or PC box, a conserved region responsible for incorporation of CBX to PRC1. Seeking to further pinpoint the interaction surfaces, we noticed a weak U2AF ligand motif (ULM) candidate on CBX7 at Pc-box proximal region [151]. We will also mutagenize the critical Trp residue to test if this is critical for U2AF2 binding. If this is enough to detach CBX7 from U2AF2, then, we will check effects of this separation on CBX7-chromatin localization and target repression. Also, we will perform CLIP with U2AF2 as well as with CBX7 in the *U2AF2* knockdown cells to test whether our 70-kDa CLIP signal belongs to U2AF2.

After validation of the interaction, we tested whether depletion of U2AF2 influences chromatin binding of CBX7. We realized that cytosolic CBX7 pool diminished, and RNA-dependent mobilization of CBX7 off the chromatin is blocked upon U2AF2 silencing and (Figure 5-9). Taking together, we initially thought that U2AF2 loss strengthened CBX7-chromatin interaction, an inference that suits more to competitive mode of interaction such that U2AF2 antagonizes CBX7 localization. To test the validity of this observation, we used spliceosome inhibitor, madrasin that is proposed to block spliceosome assembly beyond A complex and performed subcellular fractionation [152]. Similarly, we observed madrasin treatment slightly decreased cytosolic pool of CBX7 (data not shown).
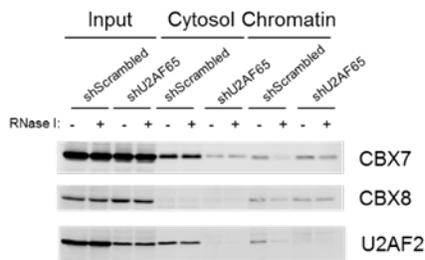
Figure 5-9: U2AF2 knockdown reduced the cytosolic pool of CBX7. Upon RNase digestion, CBX7 still remains associated with chromatin in *U2AF2* knockdown cells. This may imply U2AF2 only renders the interaction of CBX7 with chromatin RNase-sensitive; it does not alter the level of chromatin-associated CBX7.

5.1.3   Future directions and working model

Our initial goal was to dissect the targeting of CBX7, thus PRC1 complexes via RNA. However, with the current data it is also possible that RNA binding or associating with splicing-associated RBPs may not simply be for targeting CBX7-PRC1; rather, CBX7 might have direct

roles in co-transcriptional regulation such as alternative splicing. Therefore, subsequent work will be addressing the possible bidirectional functional interaction between polycomb repressive complex and the splicing control. In order to answer this question in unbiased manner, we are performing RNA-seq experiment in PC3 prostate cancer cell line, in which CBX7 was found to play oncogenic roles in RNA-dependent fashion, for determination of CBX7-dependent splicing events. [141]. Upon identification of splicing defects or alternative splicing events, we will validate these hits using RT-qPCR and RNA-immunoprecipitation for direct target search and then perform mechanistic studies to understand (1) the level and direction at which CBX7 intervenes with splicing, (2) the interplay between chromatin-binding role of CBX7 and its potential role in splicing by comparing these hits with genomic binding sites of CBX7.
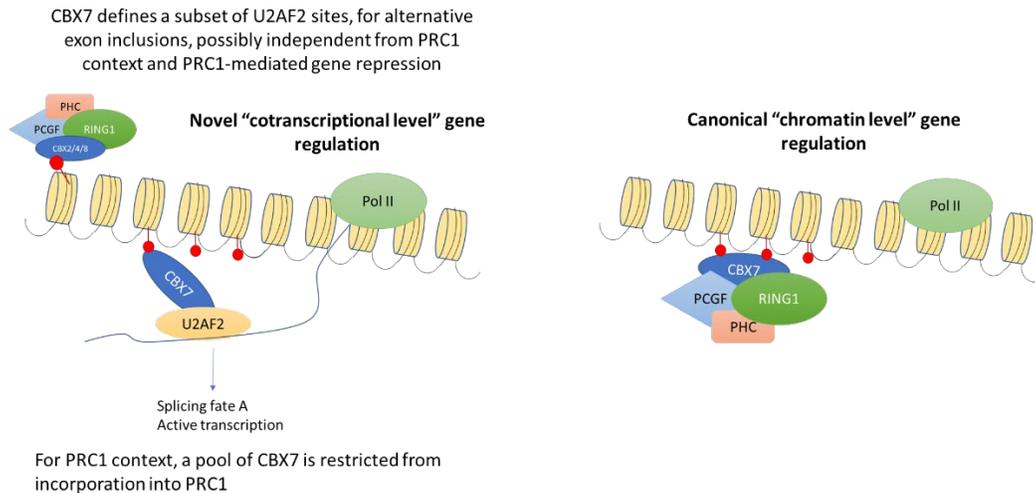
Figure 5-10 Current working model of RNA-dependent functions of CBX7. We think that a pool of CBX7 associates with splicing auxiliary factors such as U2AF2 and indirectly RNA, which is potentially independent from PRC1 context. In this case, CBX7 shuttles between cytosol and nucleus possibly as a part of mRNA export events [153]. If splicing is halted, CBX7 is retained in the nucleus, possibly incorporating with PRC1. In broader perspective, how this model goes hand in hand with transcriptional inhibition data (Figure 5-3) needs further investigation. It is possible that U2AF2 favors CBX7 export, as itself can localize to cytosol [154].

APPENDIX

SUPPLEMENTARY FIGURES FOR CHAPTER 4

Supplementary Figure 1 (Related to Figure 4-3): **a** End-point tumor images from LNCaP xenograft study on NRG mice. The graph showing the tumor masses from control and doxycycline-treated mice. Each dot represents the individual tumor mass. Two-tailed t test with Welch's correction yielded p=0.0018 (**). **b** dBRD9 reduced the growth of LNCaP, VCaP and 22Rv1 cells. The cells were plated with or without 0.5 μM dBRD9 treatment. The compound was refreshed in every two days and the cells were counted 6-10 days post-plating. Cell counts were normalized to DMSO control. The graphs depicted mean of two biological replicates each with >3 technical replicates, and error bars represent the standard deviation from normalized counts of biological replicates. Immunoblotting depicting BRD9 levels in LNCaP and VCaP cells upon treatment with a series of dBRD9 concentration for 5 days. **c** HEK293T or **d** RWPE-1 normal prostate cell line was transduced with *BRD9* shRNA vectors and selected with puromycin. The cells were counted after 4-6 days post-plating once the control cells reached confluency (n=2 independent experiments with >3 technical replicates each; error bars represent sd). Unlike LNCaP, VCaP or 22Rv1, HEK293T and RWPE-1 the cells were maintained viability at least 20 days-post transductions without any sign of evident growth defects. Knockdown was validated twice using nuclear extracts.

a

Tumor weight

b

LNCaP   22Rv1   VCaP
p=0.0171   p=0.0456   p=0.0001

c

shScramble   shBRD9_34   shBRD9_81

HEK293T

HEK293T

d

RWPE-1

RWPE-1
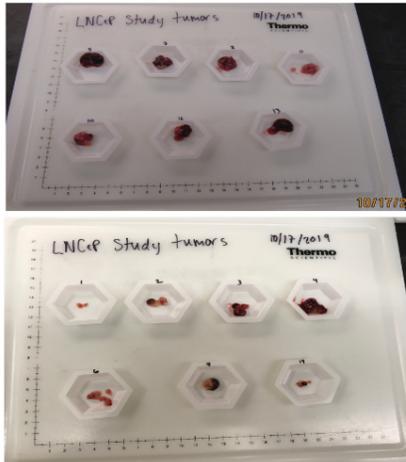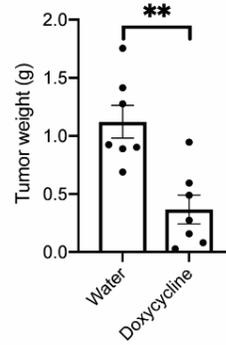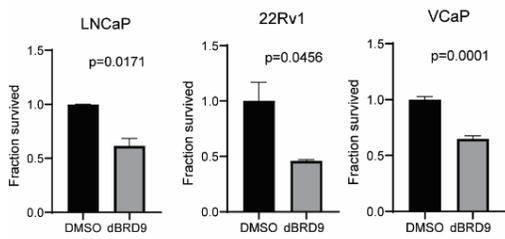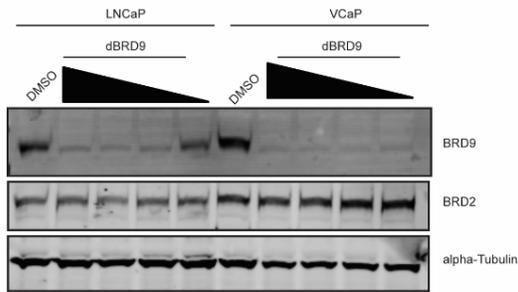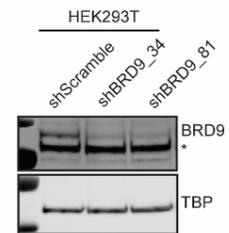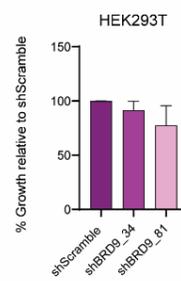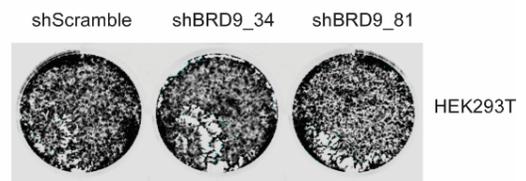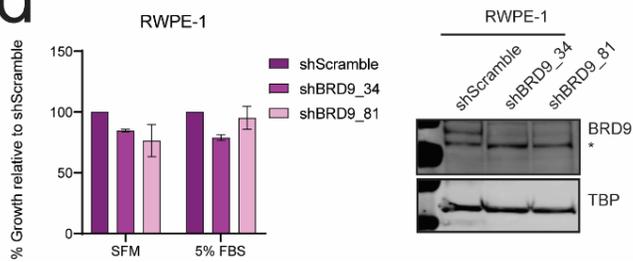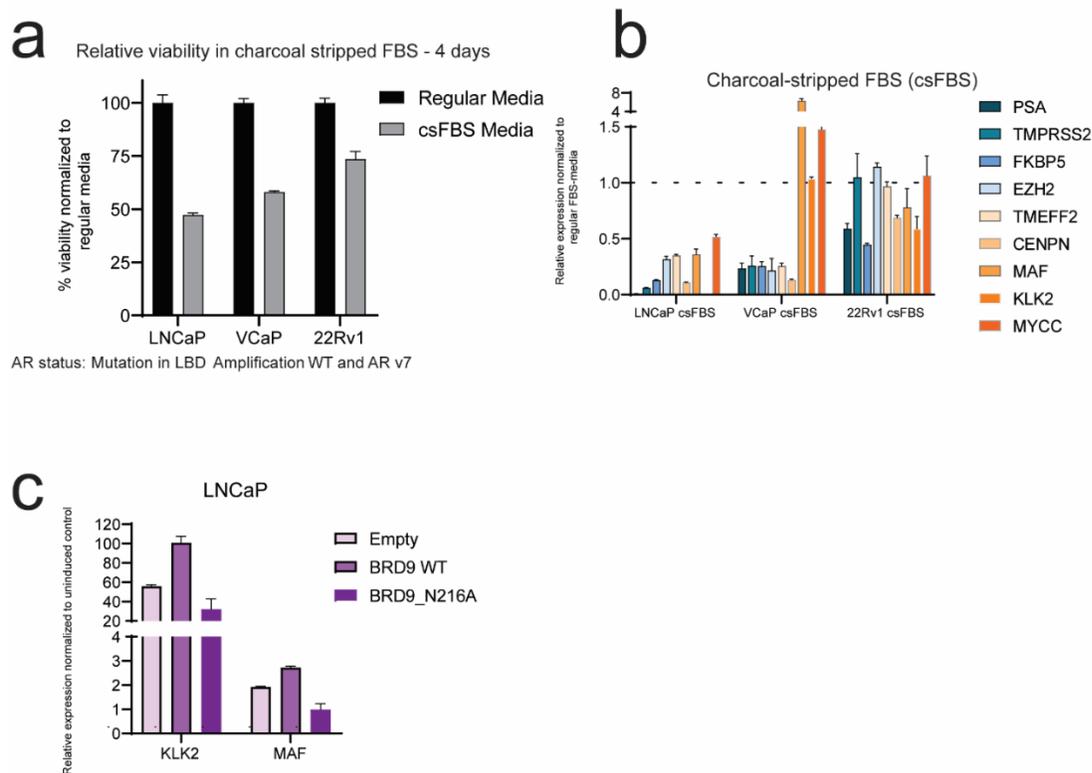
Supplementary Figure 2 (Related to Figure 4-4.): **a** Influence of charcoal-stripped FBS containing media on proliferation of AR-positive prostate cancer cell lines LNCaP, VCaP and 22Rv1. The AR status was depicted for each cell line under viability plot. The cells were plated on regular media or csFBS-containing media for 4 days. The viability was assessed with CellTiter Glo assay The data was reflective of multiple observations, represents one biological repeat with n=4 technical repeats. Percent viability was expressed relative to the signal from the cells growing on regular media. Error bars represent sd. **b I**nfluence of charcoal-stripped FBS containing media on expression of AR targets in LNCaP, VCaP and 22Rv1. LNCaP cells are more sensitive to androgen deprivation in terms of the extend of gene downregulation and the number of downregulated genes. Certain genes were commonly downregulated while others have more cell type-specific behavior. Note that 22Rv1 expresses wildtype as well as variant AR that lacks ligand binding domain (thus, this variant is inherently androgen-insensitive), as a reasoning for resistance to androgen deprivation in terms of viability and gene expression. **c** LNCaP cells with wild type BRD9 leads to higher level of AR-target gene induction. LNCaP cells were induced with doxycycline (20 ng/mL) to express wild type BRD9 (BRD-WT) or bromodomain-mutant BRD9 (BRD9 N216A) in charcoal-stripped FBS containing medium (csFBS) for two days. Cells were cultured in csFBS containing medium or regular FBS containing medium for additional 8 hours and RNA was harvested. RT-qPCR was performed for AR target genes and the relative expression values of each cell line (Empty, BRD9-WT or BRD9_N216A) were normalized to uninduced control of each group. Results were representative of two independent experiments; error bars depict standard deviation of technical replicates (n=3).
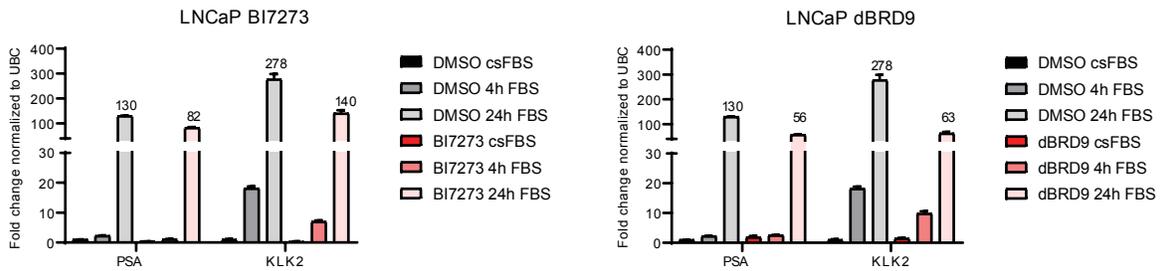
Supplementary Figure 3 (Related to Figure 4-4): Androgen-induced gene expression is retarded with another BRD9 probe (BI7273) and BRD9 degrader (dBRD9). LNCaP cells were plated on charcoal-stripped FBS containing media (csFBS) with 0.5 μM dBRD9, 10 μM BI7273 or DMSO for two days. Androgen-dependent gene expression was induced for 4 and 24 hours by switching to regular FBS containing media (FBS). RNA was harvested at the end of the treatment periods. For the RT-qPCR, the relative expression was normalized to DMSO csFBS for each sample. Note that the same DMSO treated samples were used for both dBRD9 and BI7273 sets. The graph represents one biological sample with three technical replicates. Error bars represents standard deviations and the numbers at the top of bars represent the mean of technical replicates.
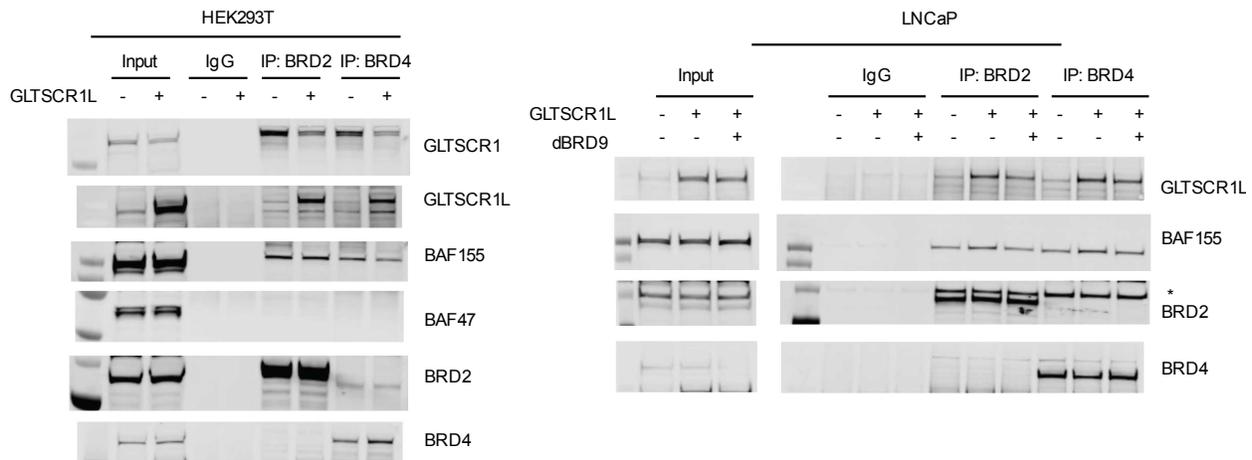
Supplementary Figure 4 (Related to Figure 4-6): *left* GLTSCR1L, once overexpressed, interacts with BRD2 and BRD4, competing with endogenous GLTSCR1. *right* This interaction is partially BRD9 dependent. LNCaP cells were induced to expressed GLTSCR1L 24 hours before 24h-treatment with 0.5 μM dBRD9. Cells were harvested and BRD2 and BRD4 were immunoprecipitated from control cells, GLTSCR1L expressing cells and GLTSCR1L expressing/dBRD9 treated cells.
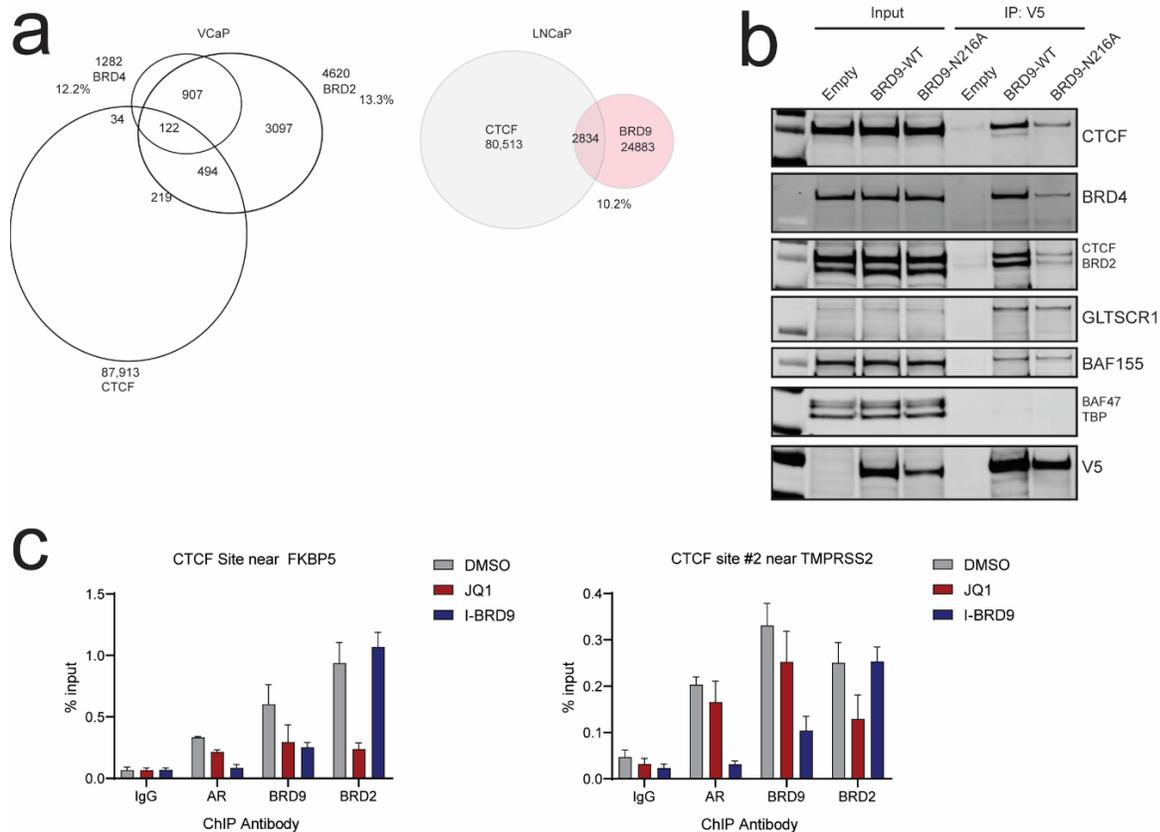
Supplementary Figure 5: BRD9 might localize CTCF sites in BET protein-dependent manner **a** Venn diagram (left) showing the overlaps between BRD2, BRD4 [68] and CTCF (ENCODE) binding sites on VCaP cells. Venn diagram (right) for comparison of CTCF (ENCODE) and BRD9 binding sites in LNCaP cells **b** CTCF association of BRD9 is bromodomain-dependent. The experiment was performed as explained in Figure 4-5b. **c** CTCF localization of BRD9 is possibly mediated through BET proteins, particularly BRD2. Genomic sites with co-enrichment of CTCF (overlapping peaks from LNCaP and VCaP CTCF ChIP-seq from ENCODE), BRD2/4, AR (VCaP from Asangani *et al.*, 2014) and BAF155 [122] were selected. LNCaP cells were grown in charcoal stripped FBS (csFBS) containing media for 24 hours. The cells were cultured for additional 20 hours in fresh csFBS-containing media with 0.5 μM JQ1 or 10 μM I-BRD9. The cells were switched to regular FBS containing media with fresh compounds and incubated for 4 hours before crosslinking and harvesting. AR, BRD9 and BRD2 were immunoprecipitated and associated DNA was purified. qPCR was performed for the sites mentioned. The enrichment was expressed relative to input and the graphs were representative of two independent experiments with 3 technical replicates. Error bars represent standard deviation.
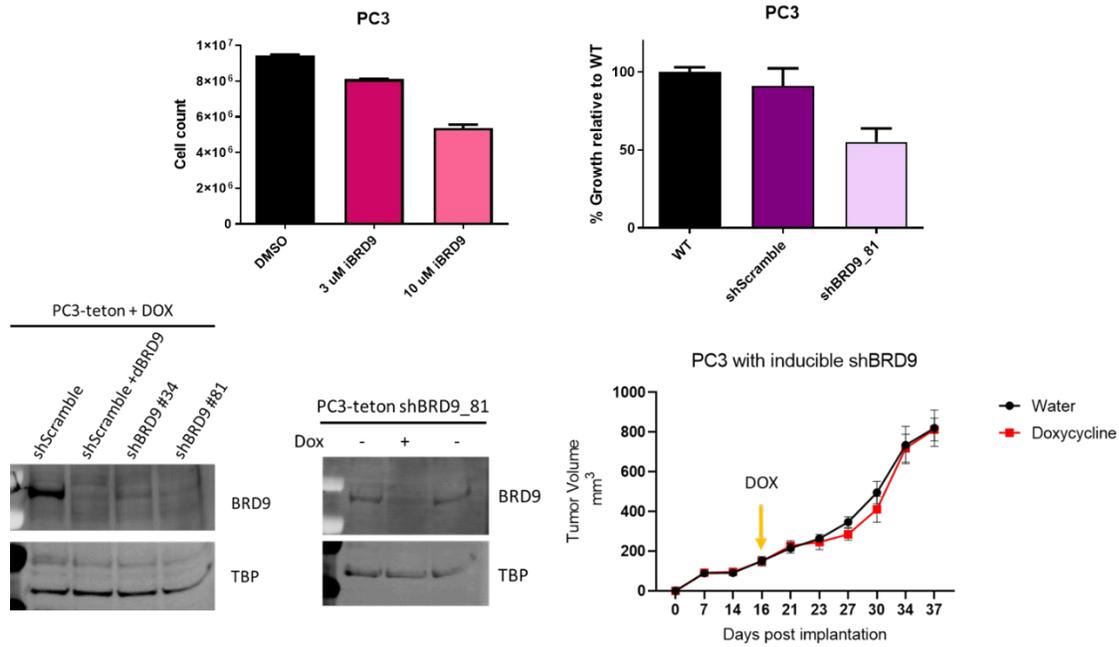
Supplementary Figure 6: BRD9 knockdown reduced the AR-negative PC3 growth in 2D culture but not in mouse xenografts.

REFERENCES

1.   Rosa, S. & Shaw, P. Insights into chromatin structure and dynamics in plants. *Biology (Basel).* **2**, 1378–410 (2013).

2.   Ma, Y. *et al.* DNA CpG hypomethylation induces heterochromatin reorganization involving the histone variant macroH2A. *J. Cell Sci.* **118**, 1607–1616 (2005).

3.   Dulac, C. Brain function and chromatin plasticity. *Nature* **465**, 728–735 (2010).

4.   Jones, P. A. Functions of DNA methylation: Islands, start sites, gene bodies and beyond. *Nat. Rev. Genet.* **13**, 484–492 (2012).

5.   Valencia, A. M. & Kadoch, C. Chromatin regulatory mechanisms and therapeutic opportunities in cancer. *Nat. Cell Biol.* **21**, 152–161 (2019).

6.   Rousseaux, S. & Khochbin, S. Histone Acylation beyond Acetylation: Terra Incognita in Chromatin Biology. *Cell J.* **17**, 1–6 (2015).

7.   Wang, Y. *et al.* KAT2A coupled with the α-KGDH complex acts as a histone H3 succinyltransferase. *Nature* **552**, 273–277 (2017).

8.   Farrelly, L. A. *et al.* Histone serotonylation is a permissive modification that enhances TFIID binding to H3K4me3. *Nature* **567**, 535–539 (2019).

9.   Schwämmle, V. *et al.* Systems level analysis of histone H3 post-translational modifications reveals features of PTM crosstalk in chromatin regulation. *Mol. Cell. Proteomics* 1–31 (2016). doi:10.1074/mcp.M115.054460

10.   Clapier, C. R., Iwasa, J., Cairns, B. R. & Peterson, C. L. Mechanisms of action and regulation of ATP-dependent chromatin-remodelling complexes. *Nat. Rev. Mol. Cell Biol.* **18**, 407–422 (2017).

11.   Hota, S. K. & Bruneau, B. G. ATP-dependent chromatin remodeling during mammalian development. *Dev.* **143**, 2882–2897 (2016).

12.   Brahma, S. *et al.* INO80 exchanges H2A.Z for H2A by translocating on DNA proximal to histone dimers. *Nat. Commun.* **8**, (2017).

13.   Kadoch, C. & Crabtree, G. R. Mammalian SWI/SNF chromatin remodeling complexes and cancer: Mechanistic insights gained from human genomics. *Sci. Adv.* **1**, e1500447–e1500447 (2015).

14.   Mashtalir, N. *et al.* Modular Organization and Assembly of SWI/SNF Family Chromatin Remodeling Complexes. *Cell* **175**, 1272-1288.e20 (2018).

15.   Alfert, A., Moreno, N. & Kerl, K. The BAF complex in development and disease. *Epigenetics and Chromatin* **12**, 1–15 (2019).

16.   Yoo, A. S., Staahl, B. T., Chen, L. & Crabtree, G. R. MicroRNA-mediated switching of chromatin-remodelling complexes in neural development. *Nature* **460**, 642–646 (2009).

17.   Sun, X. *et al.* Cardiac-enriched BAF chromatin-remodeling complex subunit Baf60c regulates gene expression programs essential for heart development and function. *Biol. Open* **7**, (2018).

18.   Lickert, H. *et al.* Baf60c is essential for function of BAF chromatin remodelling complexes in heart development. **432**, 107–112 (2004).

19.   Kadoch, C. *et al.* Proteomic and bioinformatic analysis of mammalian SWI/SNF complexes identifies extensive roles in human malignancy. *Nat. Genet.* **45**, 592–601 (2013).

20.   Nakayama, R. T. *et al.* SMARCB1 is required for widespread BAF complex–mediated activation of enhancers and bivalent promoters. *Nat. Genet.* **49**, 1613–1623 (2017).

21.   Valencia, A. M. *et al.* Recurrent SMARCB1 Mutations Reveal a Nucleosome Acidic Patch Interaction Site That Potentiates mSWI / SNF Complex Chromatin Article Recurrent SMARCB1 Mutations Reveal a Nucleosome Acidic Patch Interaction Site That Potentiates mSWI / SNF Complex Chromatin. *Cell* **179**, 1342-1356.e23 (2019).

22.   Hoffman, G. R. *et al.* Functional epigenetics approach identifies BRM/SMARCA2 as a critical synthetic lethal target in BRG1-deficient cancers. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 3128–3133 (2014).

23.   Ehrenhöfer-Wölfer, K. *et al.* SMARCA2-deficiency confers sensitivity to targeted inhibition of SMARCA4 in esophageal squamous cell carcinoma cell lines. *Sci. Rep.* **9**, 1–12 (2019).

24.   Helming, K. C. *et al.* ARID1B is a specific vulnerability in ARID1A-mutant cancers. *Nat. Med.* **20**, 251–254 (2014).

25.   Nieder, B. *et al.* Targeting ARID1A-mutant colorectal cancer : depletion of ARID1B increases radiosensitivity and modulates DNA damage response. 1–9 (2019). doi:10.1038/s41598-019-54757-z

26.   Kelso, T. W. R. *et al.* Chromatin accessibility underlies synthetic lethality of SWI/SNF subunits in ARID1A-mutant cancers. *Elife* **6**, (2017).

27.   Michel, B. C. *et al.* A non-canonical SWI/SNF complex is a synthetic lethal target in cancers driven by BAF complex perturbation. *Nat. Cell Biol.* **20**, 1410–1420 (2018).

28.   Alpsoy, A. & Dykhuizen, E. C. Glioma tumor suppressor candidate region gene 1 (GLTSCR1) and its paralog GLTSCR1-like form SWI/SNF chromatin remodeling subcomplexes. *J. Biol. Chem.* **1**, jbc.RA117.001065 (2018).

29.   Gatchalian, J. *et al.* A non-canonical BRD9-containing BAF chromatin remodeling complex regulates naive pluripotency in mouse embryonic stem cells. *Nat. Commun.* **9**, 5139 (2018).

30.   Brien, G. L. *et al.* Targeted degradation of BRD9 reverses oncogenic gene expression in synovial sarcoma. *Elife* **7**, (2018).

31.   Wang, X. *et al.* BRD9 defines a SWI/SNF sub-complex and constitutes a specific vulnerability in malignant rhabdoid tumors. *Nat. Commun.* **10**, 1–11 (2019).

32.   Middeljans, E. *et al.* SS18 Together with Animal-Specific Factors Defines Human BAF-Type SWI/SNF Complexes. *PLoS One* **7**, e33834 (2012).

33.   Hohmann, A. F. *et al.* Sensitivity and engineered resistance of myeloid leukemia cells to BRD9 inhibition. *Nat. Chem. Biol.* **12**, 672 (2016).

34.   Wei, Z. *et al.* Vitamin D Switches BAF Complexes to Protect β Cells. *Cell* **173**, 1135-1149.e15 (2018).

35.   Wang, X. *et al.* SMARCB1-mediated SWI/SNF complex function is essential for enhancer regulation. *Nat. Genet.* **49**, 289–295 (2017).

36.   Kadoch, C. & Crabtree, G. R. Reversible disruption of mSWI/SNF (BAF) complexes by the SS18-SSX oncogenic fusion in synovial sarcoma. *Cell* **153**, 71–85 (2013).

37.   Wang, W. *et al.* Purification and biochemical heterogeneity of the mammalian SWI-SNF complex. *EMBO J.* **15**, 5370–5382 (1996).

38.   Hargreaves, D. C. & Crabtree, G. R. ATP-dependent chromatin remodeling: genetics, genomics and mechanisms. *Cell Res.* **21**, 396–420 (2011).

39.   Dhalluin, C. *et al.* Structure and ligand of a histone acetyltransferase bromodomain. *Nature* **399**, 491–496 (1999).

40.   Owen, D. J. The structural basis for the recognition of acetylated histone H4 by the bromodomain of histone acetyltransferase Gcn5p. *EMBO J.* **19**, 6141–6149 (2000).

41.   Ronan, J. L., Wu, W. & Crabtree, G. R. From neural development to cognition: unexpected roles for chromatin. *Nat. Rev. Genet.* **14**, 347–359 (2013).

42.   Zinzalla, G. A New Way Forward in Cancer Drug Discovery: Inhibiting the SWI/SNF Chromatin Remodelling Complex. *ChemBioChem* **17**, 677–682 (2016).

43.   Kalpana, G., Marmon, S., Wang, W., Crabtree, G. & Goff, S. Binding and stimulation of HIV-1 integrase by a human homolog of yeast transcription factor SNF5. *Science (80-.).* **266**, 2002–2006 (1994).

44. Rafati, H. *et al.* Repressive LTR Nucleosome Positioning by the BAF Complex Is Required for HIV Latency. *PLoS Biol.* **9**, e1001206 (2011).

45. Mathies, L. D., Aliev, F., Davies, A. G., Dick, D. M. & Bettinger, J. C. Variation in SWI/SNF Chromatin Remodeling Complex Proteins is Associated with Alcohol Dependence and Antisocial Behavior in Human Populations. *Alcohol. Clin. Exp. Res.* **41**, 2033–2040 (2017).

46. Mathies, L. D. *et al.* SWI/SNF chromatin remodeling regulates alcohol response behaviors inCaenorhabditis elegansand is associated with alcohol dependence in humans. *Proc. Natl. Acad. Sci.* **112**, 3032–3037 (2015).

47. Hang, C. T. *et al.* Chromatin regulation by Brg1 underlies heart muscle development and disease. *Nature* **466**, 62–67 (2010).

48. Wu, J. I. Diverse functions of ATP-dependent chromatin remodeling complexes in development and cancer. *Acta Biochim. Biophys. Sin. (Shanghai)*. **44**, 54–69 (2011).

49. Lopes Cardoso, D. & Sharpe, C. Relating protein functional diversity to cell type number identifies genes that determine dynamic aspects of chromatin organisation as potential contributors to organismal complexity. *PLoS One* **12**, e0185409 (2017).

50. Mani, U., S., A. S., Goutham R. N., A. & Mohan S., S. SWI/SNF Infobase—An exclusive information portal for SWI/SNF remodeling complex subunits. *PLoS One* **12**, e0184445 (2017).

51. Wu, J. I., Lessard, J. & Crabtree, G. R. Understanding the Words of Chromatin Regulation. *Cell* **136**, 200–206 (2009).

52. Phelan, M. L., Sif, S., Narlikar, G. J. & Kingston, R. E. Reconstitution of a Core Chromatin Remodeling Complex from SWI/SNF Subunits. *Mol. Cell* **3**, 247–253 (1999).

53. Ho, L. & Crabtree, G. R. Chromatin remodelling during development. *Nature* **463**, 474–484 (2010).

54. Staahl, B. T. *et al.* Kinetic Analysis of npBAF to nBAF Switching Reveals Exchange of SS18 with CREST and Integration with Neural Developmental Pathways. *J. Neurosci.* **33**, 10348–10361 (2013).

55. Raab, J. R., Resnick, S. & Magnuson, T. Genome-Wide Transcriptional Regulation Mediated by Biochemically Distinct SWI/SNF Complexes. *PLOS Genet.* **11**, e1005748 (2015).

56. Mathur, R. *et al.* ARID1A loss impairs enhancer-mediated gene regulation and drives colon cancer in mice. *Nat. Genet.* **49**, 296–302 (2016).

57. Wurster, A. L. *et al.* IL-10 transcription is negatively regulated by BAF180, a component of the SWI/SNF chromatin remodeling enzyme. *BMC Immunol.* **13**, 9 (2012).

58.   Mayes, K., Qiu, Z., Alhazmi, A. & Landry, J. W. ATP-Dependent Chromatin Remodeling Complexes as Novel Targets for Cancer Therapy. *Advances in Cancer Research* 183–233 (2014). doi:10.1016/b978-0-12-800249-0.00005-6

59.   Hohmann, A. F. & Vakoc, C. R. A rationale to target the SWI/SNF complex for cancer therapy. *Trends Genet.* **30**, 356–363 (2014).

60.   Ho, L. *et al.* An embryonic stem cell chromatin remodeling complex, esBAF, is essential for embryonic stem cell self-renewal and pluripotency DEVELOPMENTAL BIOLOGY. *PNAS March* **31**, 5181–5186 (2009).

61.   Hein, M. Y. *et al.* A Human Interactome in Three Quantitative Dimensions Organized by Stoichiometries and Abundances. *Cell* **163**, 712–723 (2015).

62.   Huttlin, E. L. *et al.* Architecture of the human interactome defines protein communities and disease networks. *Nature* **545**, 505–509 (2017).

63.   Uhlen, M. *et al.* Tissue-based map of the human proteome. *Science (80-.).* **347**, 1260419 (2015).

64.   Porter, E. G. & Dykhuizen, E. C. Individual Bromodomains of Polybromo-1 Contribute to Chromatin Association and Tumor Suppression in Clear Cell Renal Carcinoma. *J. Biol. Chem.* **292**, 2601–2610 (2017).

65.   Rahman, S. *et al.* The Brd4 Extraterminal Domain Confers Transcription Activation Independent of pTEFb by Recruiting Multiple Proteins, Including NSD3. *Mol. Cell. Biol.* **31**, 2641–2652 (2011).

66.   Crowe, B. L. *et al.* Structure of the Brd4 ET domain bound to a C-terminal motif from γ-retroviral integrases reveals a conserved mechanism of interaction. *Proc. Natl. Acad. Sci.* **113**, 2086–2091 (2016).

67.   Wu, S.-Y. Y., Lee, A.-Y. Y., Lai, H.-T. T., Zhang, H. & Chiang, C.-M. M. Phospho Switch Triggers Brd4 Chromatin Binding and Activator Recruitment for Gene-Specific Targeting. *Mol. Cell* **49**, 843–857 (2013).

68.   Asangani, I. a *et al.* Therapeutic targeting of BET bromodomain proteins in castration-resistant prostate cancer. *Nature* **510**, 278–282 (2014).

69.   Shen, C. *et al.* NSD3-Short Is an Adaptor Protein that Couples BRD4 to the CHD8 Chromatin Remodeler. *Mol. Cell* **60**, 847–859 (2015).

70.   Dickinson, M. E. *et al.* High-throughput discovery of novel developmental phenotypes. *Nature* **537**, 508–514 (2016).

71.   Forbes, S. A. *et al.* COSMIC: mining complete cancer genomes in the Catalogue of Somatic Mutations in Cancer. *Nucleic Acids Res.* **39**, D945–D950 (2010).

72. Sun, A. *et al.* Aberrant expression of SWI/SNF catalytic subunits BRG1/BRM is associated with tumor development and increased invasiveness in prostate cancers. *Prostate* **67**, 203–213 (2006).

73. Wang, X. *et al.* Oncogenesis Caused by Loss of the SNF5 Tumor Suppressor Is Dependent on Activity of BRG1, the ATPase of the SWI/SNF Chromatin Remodeling Complex. *Cancer Res.* **69**, 8094–8101 (2009).

74. Krämer, K., Moreno, N., Frühwald, M. & Kerl, K. BRD9 Inhibition, Alone or in Combination with Cytostatic Compounds as a Therapeutic Approach in Rhabdoid Tumors. *Int. J. Mol. Sci.* **18**, 1537 (2017).

75. Shi, J. *et al.* Role of SWI/SNF in acute leukemia maintenance and enhancer-mediated Myc regulation. *Genes Dev.* **27**, 2648–2662 (2013).

76. Martin, L. J. *et al.* Structure-Based Design of an in Vivo Active Selective BRD9 Inhibitor. *J. Med. Chem.* **59**, 4462–4475 (2016).

77. Conrad, R. J. *et al.* The Short Isoform of BRD4 Promotes HIV-1 Latency by Engaging Repressive SWI / SNF Chromatin- Remodeling Complexes Article The Short Isoform of BRD4 Promotes HIV-1 Latency by Engaging Repressive SWI / SNF Chromatin- Remodeling Complexes. *Mol. Cell* **67**, 1001-1012.e6 (2017).

78. Zuber, J. *et al.* RNAi screen identifies Brd4 as a therapeutic target in acute myeloid leukaemia. *Nature* **478**, 524–528 (2011).

79. Shi, J. *et al.* Discovery of cancer drug targets by CRISPR-Cas9 screening of protein domains. *Nat. Biotechnol.* **33**, 1–10 (2015).

80. Porter, E. G., Connelly, K. E. & Dykhuizen, E. C. Sequential Salt Extractions for the Analysis of Bulk Chromatin Binding Properties of Chromatin Modifying Complexes. *J. Vis. Exp.* (2017). doi:10.3791/55369

81. Stanton, B. Z., Hodges, C., Crabtree, G. R. & Zhao, K. A General Non-Radioactive ATPase Assay for Chromatin Remodeling Complexes. *Curr. Protoc. Chem. Biol.* **9**, 1–10 (2017).

82. Wang, T., Wei, J. J., Sabatini, D. M. & Lander, E. S. Genetic Screens in Human Cells Using the CRISPR-Cas9 System. *Science (80-.).* **343**, 80–84 (2013).

83. Ran, F. A. *et al.* Genome engineering using the CRISPR-Cas9 system. *Nat. Protoc.* **8**, 2281–308 (2013).

84. Lessard, J. *et al.* An Essential Switch in Subunit Composition of a Chromatin Remodeling Complex during Neural Development. *Neuron* **55**, 201–215 (2007).

85. Aizawa, H. *et al.* Dendrite Development Regulated by CREST, a Calcium-Regulated Transcriptional Activator. *Science (80-.).* **303**, 1634–1640 (2004).

86. Priam, P. *et al.* SMARCD2 subunit of SWI/SNF chromatin-remodeling complexes mediates granulopoiesis through a CEBPϵ dependent mechanism. *Nat. Genet.* **49**, 753–764 (2017).

87. He, L. *et al.* BAF200 is required for heart morphogenesis and coronary artery development. *PLoS One* **9**, 1–8 (2014).

88. Wang, Z. *et al.* Polybromo protein BAF180 functions in mammalian cardiac chamber maturation. *Genes Dev.* **18**, 3106–3116 (2004).

89. Kim, Y., Andrés Salazar Hernández, M., Herrema, H., Delibasi, T. & Park, S. W. The role of BRD7 in embryo development and glucose metabolism. *J. Cell. Mol. Med.* **20**, 1561–1570 (2016).

90. Hansen, J. *et al.* A large-scale, gene-driven mutagenesis approach for the functional analysis of the mouse genome. *Proc. Natl. Acad. Sci. U. S. A.* **100**, 9918–9922 (2003).

91. Soni, S. *et al.* Absence of Erythroblast Macrophage Protein (Emp) Leads to Failure of Erythroblast Nuclear Extrusion. *J. Biol. Chem.* **281**, 20181–20189 (2006).

92. McDevitt, M. A., Shivdasani, R. A., Fujiwara, Y., Yang, H. & Orkin, S. H. A 'knockdown' mutation created by cis-element gene targeting reveals the dependence of erythroid cell maturation on the level of transcription factor GATA-1. *Proc. Natl. Acad. Sci. U. S. A.* **94**, 6781–6785 (1997).

93. Kusakabe, M. *et al.* c-Maf plays a crucial role for the definitive erythropoiesis that accompanies erythroblastic island formation in the fetal liver. *Blood* **118**, 1374–1385 (2011).

94. Palis, J. Primitive and definitive erythropoiesis in mammals. *Front. Physiol.* **5 JAN**, 1–9 (2014).

95. McGrath, K. E. *et al.* A transient definitive erythroid lineage with unique regulation of the β-globin locus in the mammalian embryo. *Blood* **117**, 4600–4608 (2011).

96. Palis, J., Malik, J., McGrath, K. E. & Kingsley, P. D. Primitive erythropoiesis in the mammalian embryo. *Int. J. Dev. Biol.* **54**, 1011–1018 (2010).

97. Dzierzak, E. & Philipsen, S. Erythropoiesis: Development and differentiation. *Cold Spring Harb. Perspect. Med.* **3**, 1–16 (2013).

98. Wang, L. *et al.* Deletion of Stk40 impairs definitive erythropoiesis in the mouse fetal liver. *Cell Death Dis.* **8**, (2017).

99. Marine, J. C. *et al.* SOCS3 is essential in the regulation of fetal liver erythropoiesis. *Cell* **98**, 617–627 (1999).

100. Sui, Z. *et al.* Tropomodulin3-null mice are embryonic lethal with anemia due to impaired erythroid terminal differentiation in the fetal liver. *Blood* **123**, 758–767 (2014).

101. Basu, P. *et al.* KLF2 is essential for primitive erythropoiesis and regulates the human and murine embryonic β-like globin genes in vivo. *Blood* **106**, 2566–2571 (2005).

102. Seu, K. G. *et al.* Unraveling macrophage heterogeneity in erythroblastic islands. *Front. Immunol.* **8**, (2017).

103. Chow, A. *et al.* CD169 + macrophages provide a niche promoting erythropoiesis under homeostasis and stress. *Nat. Med.* **19**, 429–436 (2013).

104. Liu, S. *et al.* A comprehensive analysis of coregulator recruitment, androgen receptor function and gene expression in prostate cancer. *Elife* **6**, (2017).

105. Li, X. *et al.* BRD4 Promotes DNA Repair and Mediates the Formation of TMPRSS2-ERG Gene Rearrangements in Prostate Cancer. *Cell Rep.* **22**, 796–808 (2018).

106. Welti, J. *et al.* Targeting Bromodomain and Extra-Terminal (BET) family proteins in Castration-Resistant Prostate Cancer (CRPC). *Clin. Cancer Res.* **24**, 3149–3162 (2018).

107. Urbanucci, A. *et al.* Androgen Receptor Deregulation Drives Bromodomain-Mediated Chromatin Alterations in Prostate Cancer. *Cell Rep.* **19**, 2045–2059 (2017).

108. Hong, C. Y., Suh, J. H., Kim, K., Gong, E. & Lee, K. Modulation of Androgen Receptor Transactivation by the SWI3-Related Gene Product (SRG3) in Multiple Ways. **25**, 4841–4852 (2005).

109. Link, K. A. *et al.* Targeting the BAF57 SWI/SNF subunit in prostate cancer: A novel platform to control androgen receptor activity. *Cancer Res.* **68**, 4551–4558 (2008).

110. Jin, M. L., Kim, Y. W. & Jeong, K. W. BAF53A regulates androgen receptor-mediated gene expression and proliferation in LNCaP cells. *Biochem. Biophys. Res. Commun.* **505**, 618–623 (2018).

111. Van De Wijngaart, D. J. *et al.* Functional screening of FxxLF-like peptide motifs identifies SMARCD1/BAF60a as an androgen receptor cofactor that modulates TMPRSS2 expression. *Mol. Endocrinol.* **23**, 1776–1786 (2009).

112. Lee, R. S. & Roberts, C. W. M. Linking the SWI/SNF complex to prostate cancer. *Nat. Genet.* **45**, 1268–1269 (2013).

113. Ding, Y. *et al.* Chromatin remodeling ATPase BRG1 and PTEN are synthetic lethal in prostate cancer. **129**, 759–773 (2019).

114. Kim, T. H. & Dekker, J. Preparation of cross-linked chromatin for chip. *Cold Spring Harb. Protoc.* **2018**, 311–313 (2018).

115. Ma, X. *et al.* High levels of glioma tumor suppressor candidate region gene 1 predicts a poor prognosis for prostate cancer. *Oncol. Lett.* **16**, 6749–6755 (2018).

116. Zhang, Y. *et al.* Analysis of the androgen receptor-regulated lncRNA landscape identifies a role for ARLNC1 in prostate cancer progression. *Nat. Genet.* **50**, 814–824 (2018).

117. De Leon, J. T. *et al.* Targeting the regulation of androgen receptor signaling by the heat shock protein 90 cochaperone FKBP52 in prostate cancer cells. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 11878–11883 (2011).

118. Heinlein, C. A. & Chang, C. Androgen receptor (AR) coregulators: An overview. *Endocr. Rev.* **23**, 175–200 (2002).

119. Wang, H. J. *et al.* KDM8/JMJD5 as a dual coactivator of AR and PKM2 integrates AR/EZH2 network and tumor metabolism in CRPC. *Oncogene* **38**, 17–32 (2019).

120. Lambert, J. P. *et al.* Interactome Rewiring Following Pharmacological Targeting of BET Bromodomains. *Mol. Cell* **73**, 621-638.e17 (2019).

121. Schick, S. *et al.* Systematic characterization of BAF mutations provides insights into intracomplex synthetic lethalities in human cancers. *Nat. Genet.* (2019). doi:10.1038/s41588-019-0477-9

122. Sandoval, G. J. *et al.* Binding of TMPRSS2-ERG to BAF Chromatin Remodeling Complexes Mediates Prostate Oncogenesis. *Mol. Cell* **71**, 554-566.e7 (2018).

123. Lee, J.-W. *et al.* RUNX3 regulates cell cycle-dependent chromatin dynamics by functioning as a pioneer factor of the restriction-point. *Nat. Commun.* **10**, 1897 (2019).

124. Denis, G. V *et al.* Identification of Transcription Complexes that Contain the Double Bromodomain Protein Brd2 and Chromatin Remodeling Machines research articles. 502–511 (2006). doi:10.1021/pr050430u

125. Wai, D. C. C. *et al.* The BRD3 ET domain recognizes a short peptide motif through a mechanism that is conserved across chromatin remodelers and transcriptional regulators. *J. Biol. Chem.* **293**, 7160–7175 (2018).

126. Wang, Q. *et al.* A Hierarchical Network of Transcription Factors Governs Androgen Receptor-Dependent Prostate Cancer Growth. *Mol. Cell* **27**, 380–392 (2007).

127. Wang, Q., Carroll, J. S. & Brown, M. Spatial and temporal recruitment of androgen receptor and its coactivators involves chromosomal looping and polymerase tracking. *Mol. Cell* **19**, 631–642 (2005).

128. Jia, L. *et al.* Genomic androgen receptor-occupied regions with different functions, defined by histone acetylation, coregulators and transcriptional capacity. *PLoS One* **3**, (2008).

129. Taslim, C. *et al.* Integrated analysis identifies a class of androgen-responsive genes regulated by short combinatorial long-range mechanism facilitated by CTCF. *Nucleic Acids Res.* **40**, 4754–4764 (2012).

130. Hsu, S. C. *et al.* The BET Protein BRD2 Cooperates with CTCF to Enforce Transcriptional and Architectural Boundaries. *Mol. Cell* **66**, 102-116.e7 (2017).

131. Cheung, K. L. *et al.* Distinct Roles of Brd2 and Brd4 in Potentiating the Transcriptional Program for Th17 Cell Differentiation. *Mol. Cell* 1–13 (2017). doi:10.1016/j.molcel.2016.12.022

132. Wu, X., Johansen, J. V. & Helin, K. Fbxl10/Kdm2b recruits polycomb repressive complex 1 to CpG islands and regulates H2A ubiquitylation. *Mol. Cell* **49**, 1134–46 (2013).

133. Tavares, L. *et al.* RYBP-PRC1 Complexes Mediate H2A Ubiquitylation at Polycomb Target Sites Independently of PRC2 and H3K27me3. *Cell* **148**, 664–678 (2012).

134. Sunwoo, H. *et al.* Polycomb Repressive Complex 1 Generates Discrete Compacted Domains that Change during Differentiation. *Mol. Cell* **65**, 432-445.e6 (2017).

135. Kaustov, L. *et al.* Recognition and Specificity Determinants of the Human Cbx Chromodomains. *J. Biol. Chem.* **286**, 521–529 (2011).

136. Morey, L. *et al.* Nonoverlapping Functions of the Polycomb Group Cbx Family of Proteins in Embryonic Stem Cells. *Cell Stem Cell* **10**, 47–62 (2012).

137. Pemberton, H. *et al.* Genome-wide co-localization of Polycomb orthologs and their effects on gene expression in human fibroblasts. *Genome Biol.* **15**, R23 (2014).

138. Vincenz, C. & Kerppola, T. K. Different polycomb group CBX family proteins associate with distinct regions of chromatin using nonhomologous protein sequences. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 16572–16577 (2008).

139. Akhtar, a, Zink, D. & Becker, P. B. Chromodomains are protein-RNA interaction modules. *Nature* **407**, 405–9 (2000).

140. Bernstein, E. *et al.* Mouse Polycomb Proteins Bind Differentially to Methylated Histone H3 and RNA and Are Enriched in Facultative Heterochromatin. *Mol. Cell. Biol.* **26**, 2560–2569 (2006).

141. Yap, K. L. *et al.* Molecular interplay of the noncoding RNA ANRIL and methylated histone H3 lysine 27 by polycomb CBX7 in transcriptional silencing of INK4a. *Mol. Cell* **38**, 662–74 (2010).

142. El Messaoudi-Aubert, S. *et al.* Role for the MOV10 RNA helicase in polycomb-mediated repression of the INK4a tumor suppressor. *Nat. Struct. Mol. Biol.* **17**, 862–868 (2010).

143. Fuller-Pace, F. V. The DEAD box proteins DDX5 (p68) and DDX17 (p72): Multi-tasking transcriptional regulators. *Biochim. Biophys. Acta - Gene Regul. Mech.* **1829**, 756–763 (2013).

144. Huang, W. *et al.* DDX5 and its associated lncRNA Rmrp modulate TH17 cell effector functions. *Nature* **528**, 517–522 (2015).

145. Li, H. *et al.* Erratum: RNA Helicase DDX5 Inhibits Reprogramming to Pluripotency by miRNA-Based Repression of RYBP and its PRC1-Dependent and -Independent Functions (Cell Stem Cell (2017) 20(4) (462???477) (S1934590916304581) (10.1016/j.stem.2016.12.002)). *Cell Stem Cell* **20**, 571 (2017).

146. Zhang, H. *et al.* RNA helicase DEAD box protein 5 regulates Polycomb repressive complex 2/Hox transcript antisense intergenic RNA function in hepatitis B virus infection and hepatocarcinogenesis. *Hepatology* **64**, 1033–1048 (2016).

147. Luco, R. F. *et al.* Regulation of Alternative Splicing by Histone Modifications Published by : American Association for the Advancement of Science Linked references are available on JSTOR for this article : Regulation of Alternative Splicing by Histone Modifications. **327**, 996–1000 (2010).

148. Andersson, R., Enroth, S., Rada-Iglesias, A., Wadelius, C. & Komorowski, J. Nucleosomes are well positioned in exons and carry characteristic histone modifications. *Genome Res.* **19**, 1732–1741 (2009).

149. Bianco, M. The CBX7 protein , whose expression is decreased in human carcinomas , positively regulates E-cadherin expression by interacting with the HDAC2 protein. (University of Naples Federico II, 2009).

150. Lamb, K. N. *et al.* Discovery and Characterization of a Cellular Potent Positive Allosteric Modulator of the Polycomb Repressive Complex 1 Chromodomain, CBX7. *Cell Chem. Biol.* **26**, 1365-1379.e22 (2019).

151. Loerch, S. & Kielkopf, C. L. Unmasking the U2AF homology motif family: A bona fide protein-protein interaction motif in disguise. *Rna* **22**, 1795–1807 (2016).

152. Pawellek, A. *et al.* Identification of small molecule inhibitors of pre-mRNA splicing. *J. Biol. Chem.* **289**, 34683–34698 (2014).

153. Cullen, B. R. Connections between the processing and nuclear export of mRNA: Evidence for an export license? *Proc. Natl. Acad. Sci. U. S. A.* **97**, 4–6 (2000).

154. Palangat, M. *et al.* The splicing factor U2AF1 contributes to cancer progression through a noncanonical role in translation regulation. *Genes Dev.* **33**, 482–497 (2019).

PUBLICATIONS

- Cai, B., Kim, D., Akhand, S., Sun, Y., Cassell, R.J., **Alpsoy, A**., Dykhuizen, E.C., Rijn, R.M. Van, Wendt, M.K., Krusemark, C.J., 2019. Selection of DNA-Encoded Libraries to Protein Targets within and on Living Cells. J. Am. Chem. Soc. 141, 17057–17061. https://doi.org/10.1021/jacs.9b08085

- Kulkarni, R.A., Bak, D.W., Wei, D., Bergholtz, S.E., Briney, C.A., Shrimp, J.H., **Alpsoy, A**., Thorpe, A.L., Bavari, A.E., Crooks, D.R., Levy, M., Florens, L., Washburn, M.P., Frizzell, N., Dykhuizen, E.C., Weerapana, E., Linehan, W.M., Meier, J.L., 2019. A chemoproteomic portrait of the oncometabolite fumarate. Nat. Chem. Biol. https://doi.org/10.1038/s41589-018-0217-y

- Torres-Zelada, E.F., Stephenson, R.E., **Alpsoy, A.**, Anderson, B.D., Swanson, S.K., Florens, L., Dykhuizen, E.C., Washburn, M.P., Weake, V.M., 2019. The Drosophila Dbf4 ortholog Chiffon forms a complex with Gcn5 that is necessary for histone acetylation and viability. J. Cell Sci. 132, jcs214072. https://doi.org/10.1242/jcs.214072

- Connelly, K.E., Weaver, T.M., **Alpsoy, A**., Gu, B.X., Musselman, C.A., Dykhuizen, E.C., 2018. Engagement of DNA and H3K27me3 by the CBX8 chromodomain drives chromatin association. Nucleic Acids Res. https://doi.org/10.1093/nar/gky1290

- Soto-Velasquez, M., Hayes, M.P., **Alpsoy, A**., Dykhuizen, E.C., Watts, V.J., 2018. A Novel CRISPR/Cas9-Based Cellular Model to Explore Adenylyl Cyclase and cAMP Signaling. Mol. Pharmacol. 94, 963 LP-972. https://doi.org/10.1124/mol.118.111849

- Shinde, A., Libring, S., **Alpsoy, A.**, Abdullah, A., Schaber, J.A., Solorio, L., Wendt, M.K., 2018. Autocrine Fibronectin Inhibits Breast Cancer Metastasis. Mol. Cancer Res. 16, 1579 LP-1589. https://doi.org/10.1158/1541-7786.MCR-18-0151

- **Alpsoy, A.**, Dykhuizen, E. C. (2018). Glioma tumor suppressor candidate region gene 1 (GLTSCR1) and its paralog GLTSCR1-like form SWI/SNF chromatin remodeling subcomplexes. J Biol Chem, 293(11), 3892–3903