Mechanisms and Vulnerabilities of Mutant SWI/SNF Complexes in Cancer

The Harvard community has made this article openly available. Please share how this access benefits you. Your story matters

<table>
<thead>
<tr>
<th>Citation</th>
<th>Walsh, Katherine Helming. 2015. Mechanisms and Vulnerabilities of Mutant SWI/SNF Complexes in Cancer. Doctoral dissertation, Harvard University, Graduate School of Arts &amp; Sciences.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Citable link</td>
<td><a href="http://nrs.harvard.edu/urn-3:HUL.InstRepos:17467343">http://nrs.harvard.edu/urn-3:HUL.InstRepos:17467343</a></td>
</tr>
<tr>
<td>Terms of Use</td>
<td>This article was downloaded from Harvard University’s DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at <a href="http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA">http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA</a></td>
</tr>
</tbody>
</table>
Mechanisms and Vulnerabilities of Mutant SWI/SNF Complexes in Cancer

A dissertation presented

by

Katherine Helming Walsh

to

The Division of Medical Sciences

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

in the subject of

Biological and Biomedical Sciences

Harvard University

Cambridge, Massachusetts

April 2015
Mechanisms and Vulnerabilities of Mutant SWI/SNF Complexes in Cancer

Abstract

Genes encoding subunits of the SWI/SNF chromatin remodeling complex are frequently mutated in a wide variety of cancers, but the mechanism of cancer formation following SWI/SNF subunit mutation is not known. As a result, improved understanding of how mutation of these subunits promotes oncogenesis could have broad relevance for human cancer therapy. In this dissertation, two distinct lines of investigation were pursued with the common goals of understanding the mechanisms by which SWI/SNF subunit mutation contributes to cancer and identifying vulnerabilities specific to SWI/SNF-mutant cancers. First, a genome-wide screen was used to find specific dependencies conferred by ARID1A mutation, the most frequently mutated subunit of the SWI/SNF complex in cancer. ARID1B was identified as the number one gene specifically essential for the growth of ARID1A-mutant cancers, and ARID1B loss was found to destabilize SWI/SNF and impair cell proliferation in ARID1A-mutant backgrounds. ARID1A and ARID1B were also found to be frequently co-mutated in cancer cell lines and primary samples. Furthermore, these proteins were found to share overlapping roles but also retain unique functions in a cell context dependent manner. Second, DNA methylation was studied as a putative epigenetic mechanism by which mutation of SMARCB1, a core SWI/SNF subunit, leads to genomically stable rhabdoid tumors. Reintroduction of SMARCB1 in rhabdoid tumor cell lines resulted in a decrease of DNA methylation at a subset of upregulated enhancers, identifying DNA methylation as a secondary change following SWI/SNF subunit perturbation. The results from these two
distinct lines of study reveal novel vulnerabilities of cancers harboring SWI/SNF mutation and elucidate partial mechanism by which mutation of two unique SWI/SNF subunits contributes to cancer. These findings provide new insight into potential therapeutic opportunities for the wide spectrum of SWI/SNF-mutant cancers.
Table of Contents

Title Page .................................................................................................................. i
Abstract .................................................................................................................... iii
Table of Contents ....................................................................................................... v
Acknowledgements ..................................................................................................... vii
Dedication .................................................................................................................... ix

Chapter 1: Introduction ............................................................................................. 1
SWI/SNF complexes and transcriptional regulation ..................................................... 2
  Introduction to SWI/SNF complexes ..................................................................... 2
  ARID1A and ARID1B ................................................................................................. 5
  Roles of SWI/SNF complexes ................................................................................. 7
SWI/SNF as a major tumor suppressor complex ....................................................... 9
  SWI/SNF mutations in cancer ................................................................................. 9
  SMARCB1 mutations and rhabdoid tumors ............................................................... 9
  SWI/SNF complexes and regulation of cancer-associated genes ......................... 10
Vulnerabilities of mutant SWI/SNF complexes in cancer ....................................... 11
Role of aberrant DNA methylation in cancer ............................................................ 14
References ................................................................................................................... 15

Chapter 2: ARID1B is a specific vulnerability in ARID1A-mutant cancers .......... 24
Abstract .................................................................................................................... 25
Contributions ............................................................................................................. 26
Results ....................................................................................................................... 28
  ARID1B is specifically essential for ARID1A-mutant cells .................................... 28
  ARID1B loss destabilizes SWI/SNF complexes in ARID1A-mutant cells ............. 31
  ARID1A and ARID1B mutations co-occur in cell lines and primary cancers .......... 38
  In vivo xenograft experiments ................................................................................. 39
Conclusion .................................................................................................................. 43
Methods .................................................................................................................... 45
References ................................................................................................................... 49

Chapter 3: Evaluation of the redundancy between ARID1A and ARID1B .......... 52
Abstract .................................................................................................................... 53
Contributions ............................................................................................................. 54
Introduction ............................................................................................................... 55
Results ....................................................................................................................... 56
  CRISPR-mediated deletion of ARID1B in ARID1A isogenic cell lines ................. 56
  shRNA-mediated knockdown of ARID1A and ARID1B ........................................ 63
  Evaluation of gene expression changes following ARID1A and ARID1B knockdown .............................................................. 71
Conclusion .................................................................................................................. 81
Methods .................................................................................................................... 81
References ................................................................................................................... 83
Chapter 4: SMARCB1 loss affects DNA methylation at a subset of upregulated enhancers

Abstract.........................................................................................................................87
Contributions ..................................................................................................................88
Introduction.....................................................................................................................89
Results.............................................................................................................................89
  SMARCA4 does not bind to regions of methylated DNA in MEFs .........................89
  SMARCB1 re-expression results in a decrease of DNA methylation at enhancers in rhabdoid tumor cell lines .................................................................92
  Addback of SMARCB1 in rhabdoid cell lines results in a specific demethylation of a subset of upregulated SWI/SNF-bound enhancers .................................................94
Conclusion......................................................................................................................98
Methods..........................................................................................................................99
References.......................................................................................................................100

Chapter 5: Discussion and future directions .................................................................102
The relationship between ARID1A and ARID1B .........................................................103
  Cell context is essential for the precise interaction between
  ARID1A and ARID1B .................................................................................................103
  ARID1A and ARID1B share overlapping roles but also retain unique functions ..........105
  ARID1A and ARID1B may bind to each other in certain contexts .........................106
  Implications for therapy.............................................................................................108
  Future directions.........................................................................................................109
The role of DNA methylation as an epigenetic mechanism in
  SMARCB1-deficient cancers.......................................................................................111
  Crucial role for enhancers in rhabdoid tumors.........................................................111
  DNA methylation change is a secondary effect following SMARCB1 loss ..........112
  Future directions.........................................................................................................115
Conclusion......................................................................................................................117
References.......................................................................................................................118
Acknowledgements

Many individuals have contributed to my graduate experience in numerous ways and deserve immense thanks. First and foremost, I would like to extend my sincere gratitude to my mentor, Dr. Charles Roberts, for his guidance and support in and out of the lab during my time in graduate school. Charlie has taught me to think critically and creatively, and his excitement for science and commitment to the lab created a rich environment for my graduate research. I would also like to acknowledge past and present members of the Roberts Lab. Each of you has enhanced my graduate experience in your own unique ways. I greatly appreciate the time you all spent editing drafts of papers and reports, discussing data, and just being there to share a laugh. I especially would like to thank Dr. Xiaofeng Wang for his wonderful collaboration and friendship, whose work ethic and scientific skill impress me every day. I would also like to extend my great thanks to Haley Manchester for her technical help through the bulk of my graduate school time. It was a privilege to watch Haley grow as a scientist and I feel lucky to call her my friend. I am also extremely grateful for the companionship and camaraderie from my fellow classmates, especially my roommate, Natalie German, and Lauren Barclay, who both provided the support and encouragement necessary to make it through the PhD. Natalie went above and beyond the role of roommate and became a member of my family who will always have a place at our Thanksgiving table.

I would never have embarked on a career in biology had it not been for two important educators in my life. Malvina Liebert, my high school biology teacher, first instilled in me a passion and curiosity for science that changed the trajectory of my
interests. Fr. Nicanor Austriaco, OP, PhD, my college advisor, professor, and PI, first taught me to think like a scientist and inspired me to pursue an advanced degree.

I would like to thank the members of my dissertation advisory committee: Dr. Levi Garraway, Dr. Bradley Bernstein, and Dr. Myles Brown. Their thoughtful feedback and critical evaluation of my data were extremely valuable in the progression of my project. Thank you to the BBS program administrators Kate Hodgins, Maria Bollinger, and Danny Gonzalez for all their hard work and dedication to the program. I would also like to acknowledge Dr. Connie Cepko, Dr. Thomas Michel, and the Leder Human Biology Program. This program truly enriched my graduate studies and was the highlight of my graduate school curriculum.

I would not be here today if it were not for an amazingly talented and compassionate team of individuals at the Dana Farber Cancer Institute and the Brigham and Women’s Hospital. Words cannot express my deep gratitude for all that you have done for me.

Finally, I would like to acknowledge my family for all their love and support. To my parents, Mary and Tom Helming, and to my sister Emily, thank you for unconditionally supporting me through the good times and the bad. I would not have made it through graduate school without you and I am grateful to know that you are always on my team. To my husband Ryan, I cannot thank you enough for standing by my side and encouraging me at every step of the way. I truly admire your thoughtfulness, patience, and wisdom, and I feel extremely fortunate to be travelling through life with you.
This work is dedicated to all those who gave me hope and strength during my darkest days. I will forever be grateful for your love and support.
Chapter 1

Introduction
**SWI/SNF complexes and transcriptional regulation**

*Introduction to SWI/SNF complexes*

A key mechanism of gene regulation occurs through the condensation and compaction of DNA around nucleosomes, the fundamental unit of chromatin\(^1\). This process allows for the DNA to fit into the nucleus of the cell and also contributes to gene expression control. The nucleosome is a histone octamer containing two molecules each of the histones H2A, H2B, H3, and H4, and this structure is wrapped in 147 base pairs of double stranded DNA. Adjacent nucleosomes are connected by linker DNA, and the coiling of nucleosomes results in higher order chromatin structures. The degree of chromatin compaction is one mechanism of transcriptional regulation, as DNA in tightly wound chromatin is virtually inaccessible to transcriptional machinery. Proteins that modify the structure of chromatin, thereby modulating DNA accessibility, are essential for the regulation of gene expression. Two classes of chromatin remodeling complexes exist: nucleosome remodeling complexes and complexes that covalently modify histones and DNA\(^1\).

SWI/SNF complexes are evolutionarily conserved multi-subunit complexes that utilize the energy of ATP hydrolysis to mobilize nucleosomes and remodel chromatin\(^2,3\). These complexes were first identified through two genetic screens in *Saccharomyces cerevisiae*. The first screen identified genes required for proper sucrose metabolism, known as sucrose non-fermenting (SNF) mutants\(^4\), and the second identified genes necessary for mating-type switching (SWI)\(^5,6\). These SNF and SWI genes were subsequently found to encode proteins that physically interact in a multi-subunit complex\(^7\) which possesses ATP-dependent nucleosome remodeling activity\(^8-10\). Yeast
SWI/SNF is reported to contain 9-12 subunits and have a molecular mass of approximately 1.14 MDa\textsuperscript{7,9}. The structure of SWI/SNF is conserved through evolution, though some subunits are unique to human SWI/SNF, such as SMARCE1, PHF10, and DPF1,2,3\textsuperscript{11}, indicating a diversified function of mammalian SWI/SNF.

Human SWI/SNF are approximately 2 MDa complexes composed of 12-15 subunits\textsuperscript{12}. These complexes contain one of the two catalytic ATPase subunits, SMARCA4 (BRG1) or SMARCA2 (BRM,) several core subunits including SMARCB1 (SNF5/INI1/BAF47) and SMARCC1 (BAF155) that are present in all SWI/SNF complexes, as well as subunits present in only some variants such as ARID1A and ARID1B, mutually exclusive subunits for BAF (BRG1-associated factor) varieties of the complexes, and PBRM1 and ARID2, specific for PBAF (polybromo BRG1-associated factor) varieties of the complexes (Table 1-1)\textsuperscript{12-19}. SMARCA4 and SMARCA2 alone are capable of mobilizing nucleosomes \textit{in vitro}, though their remodeling activity is increased when combined with the other core subunits\textsuperscript{3}. Recent studies have identified novel variant subunits, such as SS18\textsuperscript{20}, suggesting that the variety of SWI/SNF complexes is more complicated than previously understood. Mammalian SWI/SNF was initially thought to be classified into two major subclasses, BAF and PBAF\textsuperscript{21}. More recently, however, due to the vast diversity of SWI/SNF subunits, it is hypothesized that hundreds of SWI/SNF variants exist, each with a conserved core but with a unique combination of alternate subunits\textsuperscript{13}. 
Table 1-1. Official gene names of SWI/SNF subunits, common aliases, and protein sizes.

List of some of the cancers in which these genes are found to be mutated.

<table>
<thead>
<tr>
<th>Official Gene Name</th>
<th>Other Names</th>
<th>Size (kDa)</th>
<th>Cancer Type</th>
<th>Subunits</th>
</tr>
</thead>
<tbody>
<tr>
<td>SMARCA4</td>
<td>BRG1</td>
<td>184.5</td>
<td>lung, ovarian, medulloblastoma, Burkitt lymphoma, colorectal, pancreatic, melanoma</td>
<td>catalytic ATPase subunits</td>
</tr>
<tr>
<td>SMARCA2</td>
<td>BRM</td>
<td>181</td>
<td>lung, colorectal, breast</td>
<td>core subunits</td>
</tr>
<tr>
<td>SMARCB1</td>
<td>SNF5, BAF47, INI1</td>
<td>44</td>
<td>rhabdoid tumor, familial schwannomatosis</td>
<td>variant subunits</td>
</tr>
<tr>
<td>SMARCC1</td>
<td>BAF155</td>
<td>155</td>
<td>colon</td>
<td>variant subunits</td>
</tr>
<tr>
<td>SMARCC2</td>
<td>BAF170</td>
<td>170</td>
<td></td>
<td>variant subunits</td>
</tr>
<tr>
<td>ARID1A</td>
<td>BAF250A</td>
<td>242</td>
<td>ovarian, hepatocellular, bladder, gastric, endometrioid, pancreatic, colon, lung, neuroblastoma, Burkitt lymphoma, melanoma</td>
<td>variant subunits</td>
</tr>
<tr>
<td>ARID1B</td>
<td>BAF250B</td>
<td>236</td>
<td>melanoma, neuroblastoma, hepatocellular, pancreatic, liver, gastric, colorectal, lung</td>
<td>variant subunits</td>
</tr>
<tr>
<td>PBRM1</td>
<td>BAF180</td>
<td>193</td>
<td>renal cell carcinoma, breast, gastric, pancreatic, lung</td>
<td>variant subunits</td>
</tr>
<tr>
<td>ARID2</td>
<td>BAF200</td>
<td>197</td>
<td>melanoma, pancreatic, hepatocellular, colorectal, lung</td>
<td>variant subunits</td>
</tr>
<tr>
<td>SMARCE1</td>
<td>BAF57</td>
<td>47</td>
<td>spinal meningioma</td>
<td>variant subunits</td>
</tr>
<tr>
<td>BRD7</td>
<td></td>
<td>74</td>
<td>breast</td>
<td>variant subunits</td>
</tr>
<tr>
<td>PHF10</td>
<td>BAF45A</td>
<td>56</td>
<td>hepatocellular</td>
<td>newly identified subunits</td>
</tr>
<tr>
<td>DPF1</td>
<td>BAF45B</td>
<td>42.5</td>
<td></td>
<td>newly identified subunits</td>
</tr>
<tr>
<td>DPF3</td>
<td>BAF45C</td>
<td>43</td>
<td></td>
<td>newly identified subunits</td>
</tr>
<tr>
<td>DPF2</td>
<td>BAF45D</td>
<td>44</td>
<td></td>
<td>newly identified subunits</td>
</tr>
<tr>
<td>ACTL6A</td>
<td>BAF53A</td>
<td>47.5</td>
<td></td>
<td>newly identified subunits</td>
</tr>
<tr>
<td>ACTL6B</td>
<td>BAF53B</td>
<td>47</td>
<td></td>
<td>newly identified subunits</td>
</tr>
<tr>
<td>SMARCD1</td>
<td>BAF66A</td>
<td>58</td>
<td></td>
<td>newly identified subunits</td>
</tr>
<tr>
<td>SMARCD2</td>
<td>BAF66B</td>
<td>59</td>
<td></td>
<td>newly identified subunits</td>
</tr>
<tr>
<td>SMARCD3</td>
<td>BAF66C</td>
<td>55</td>
<td></td>
<td>newly identified subunits</td>
</tr>
<tr>
<td>ACTB</td>
<td></td>
<td>41.5</td>
<td></td>
<td>newly identified subunits</td>
</tr>
<tr>
<td>BCL11A</td>
<td></td>
<td>91</td>
<td>breast</td>
<td>newly identified subunits</td>
</tr>
<tr>
<td>BCL11B</td>
<td></td>
<td>95.5</td>
<td>leukemia</td>
<td>newly identified subunits</td>
</tr>
<tr>
<td>BCL7A</td>
<td></td>
<td>26</td>
<td></td>
<td>newly identified subunits</td>
</tr>
<tr>
<td>BCL7B</td>
<td></td>
<td>22</td>
<td></td>
<td>newly identified subunits</td>
</tr>
<tr>
<td>BCL7C</td>
<td></td>
<td>26</td>
<td></td>
<td>newly identified subunits</td>
</tr>
<tr>
<td>BRD9</td>
<td></td>
<td>67</td>
<td>lung</td>
<td>newly identified subunits</td>
</tr>
<tr>
<td>SS18</td>
<td>SYT</td>
<td>46</td>
<td>synovial sarcoma</td>
<td>newly identified subunits</td>
</tr>
</tbody>
</table>
ARID1A and ARID1B

Of particular interest for this project are ARID1A and ARID1B, mutually exclusive members of BAF SWI/SNF complexes. Both proteins are named for their ARID (AT-rich interacting domain), a DNA binding domain that spans approximately 100 amino acid residues, with the major groove contact site being a helix-turn-helix motif. ARID1A and ARID1B are 80% identical across the ARID domain and are 66% homologous across the length of the entire protein. ARID1A also possesses glutamine rich regions and several LXXLL motifs, putative regions for nuclear hormone receptor binding, not conserved in ARID1B. ARID1B has been shown to assemble into an E3 ubiquitin ligase via a BC box binding motif, which targets H2B at lysine 120 for monoubiquitination. Both proteins have broad tissue distribution, and they are the mammalian orthologues of Osa in Drosophila melanogaster and Swi1p in Saccharomyces cerevisiae. ARID1A and ARID1B have been reported to have opposing functions in response to cell cycle arrest in a pre-osteoblast cell line model, where ARID1A plays a role in cell cycle repression and ARID1B promotes cell proliferation. During the cell cycle, ARID1A and ARID1B are expressed with varying kinetics. ARID1A accumulates at the G0 phase and then decreases throughout the cell cycle. ARID1B, on the other hand, remains present at low levels through all phases of the cell cycle, including mitosis.

The role of these proteins in embryonic stem (ES) cell pluripotency and maintenance has also been studied. ARID1A is crucial for early germ layer formation, as loss of ARID1A in early murine ES cells results in the absence of a mesodermal layer and a developmental arrest. As a result, Arid1a knockout in mice is embryonic lethal.
ARID1A deficiency can also negatively affect the pluripotency of ES cells, and ARID1A has been shown to affect the expression of important self-renewal genes, including Sox2, Utf1, and Oct4. Similarly, ARID1B has an important role in ES cells, as ARID1B has been shown to be required for the proper maintenance of pluripotency and proliferation of mouse ES cells.

Studies have characterized ARID1A as a *bona fide* tumor suppressor. Re-expression of wild type ARID1A in ARID1A-mutant ovarian cancer cells abrogates cell proliferation and xenograft tumor growth in mice. Silencing of ARID1A via RNAi in non-transformed epithelial cells increases cell proliferation and tumorigenicity. In addition, known p53 target genes CDKN1A and SMAD3 were identified as downstream targets of ARID1A. Furthermore, NIH3T3 cells can be transformed by ARID1A depletion, as treatment of cells with antisense ARID1A results in anchorage independent growth. A third study demonstrated that in gastric cell lines expressing ARID1A, knockdown of ARID1A with shRNA and siRNA increased cellular proliferation. Similarly, re-expression of ARID1A in ARID1A-deficient gastric cell lines decreased cell proliferation. Taken together, these experiments support the conclusion that ARID1A is a validated tumor suppressor gene.

A unique role for ARID1B in intellectual disabilities has been described. The gene encoding this protein is frequently mutated in diseases impacting intellectual abilities, such as Coffin-Siris syndrome (CSS) and Nicolaides-Baraitser syndrome (NCBRS). ARID1B is highly expressed in the brain and it plays an important role in the development and function of the human brain. Haploinsufficiency of ARID1B is associated with corpus callosum abnormalities, speech impairment, and autism. These
findings establish a role for ARID1B, distinct from ARID1A, in proper neural development.

Roles of SWI/SNF complexes

SWI/SNF complexes interact with transcription factors, co-activators and co-repressors, and are capable of mobilizing nucleosomes at target promoters and enhancers to modulate gene expression (Figure 1-1)\textsuperscript{14,38–40}. Additionally, they have been implicated in various types of DNA repair\textsuperscript{41–45}. These complexes have also been shown to serve roles in the transcriptional regulation of lineage specification and development in numerous model systems. For example, SWI/SNF complexes contribute to the development of T cells\textsuperscript{46,47}, hepatocytes\textsuperscript{48}, oligodendrocytes\textsuperscript{40}, and the maintenance of embryonic stem cell self-renewal and pluripotency\textsuperscript{28,49}. SWI/SNF complexes have also been implicated in melanocyte differentiation via MITF-mediated recruitment of SWI/SNF to melanocyte-specific promoters followed by subsequent structural chromatin changes and gene activation\textsuperscript{50}. Specificity of SWI/SNF complexes in the control of these developmental programs is achieved in part through restricted expression and combinatorial assembly of variant SWI/SNF subunits. The SMARCD3 (BAF60C) subunit is expressed specifically in the embryonic heart, where it is essential for the control of cardiac development\textsuperscript{51}. Similarly, a switch from the PHF10 (BAF45A) and ACTL6A (BAF53A) subunits, which are expressed in neural stem cells, to DPF1 (BAF45B), DPF3 (BAF45C), and ACTL6B (BAF53B) subunits is essential to control the transition of neural progenitors into post-mitotic mature neurons\textsuperscript{52,53}. Such switching can modulate interaction with specific transcription factors\textsuperscript{54} and facilitates differential
activation of transcriptional pathways. Ultimately, via combinatorial inclusion of variant subunits, several hundred versions of SWI/SNF complexes may exist\textsuperscript{13} and serve instructive roles in the control of fate specification.

\begin{figure}
\centering
\includegraphics[width=\textwidth]{mSWI-SNF.png}
\caption{	extbf{Mammalian SWI/SNF complex (mBAF).} SWI/SNF complexes are found in two major subtypes, BAF and PBAF, and are comprised of multiple subunits (top). These complexes modulate transcription by mobilizing nucleosomes and by recruiting transcription factors to the DNA. Subunits found mutated in cancer are denoted by a red star and are described in Table 1-1.}
\end{figure}

\textbf{Figure 1-1. SWI/SNF complexes modulate transcription.} SWI/SNF complexes are found in two major subtypes, BAF and PBAF, and are comprised of multiple subunits (top). These complexes modulate transcription by mobilizing nucleosomes and by recruiting transcription factors to the DNA. Subunits found mutated in cancer are denoted by a red star and are described in Table 1-1.
SWI/SNF as a major tumor suppressor complex

SWI/SNF mutations in cancer

Genes encoding many subunits of SWI/SNF complexes have been found to be recurrently mutated in 20% of all human cancers. Multiple subunits are mutated in a wide spectrum of cancers, emphasizing the concept that individual subunits are important in varying cell types and lineages (Figure 1-1 and Table 1-1). ARID1A has emerged as the most frequently mutated subunit of SWI/SNF complexes, first noted for its mutation status in ovarian clear cell carcinoma. ARID1B mutations have also been identified in cancers such as neuroblastoma and pancreatic cancer. SWI/SNF is broadly mutated in cancer, at a frequency similar to TP53 mutation, underscoring the importance of this complex in tumor suppression. As a result, an understanding of the mechanism by which SWI/SNF regulates transcription and the effect of mutations on proper SWI/SNF function will be important for improved cancer therapy.

SMARCB1 mutations and rhabdoid tumors

The first clue linking SWI/SNF complexes to cancer came in the late 1990s when mutations of the gene encoding the SMARCB1 (SNF5/INI1/BAF47) subunit were identified in rhabdoid tumors (RT), a rare but highly aggressive type of cancer that strikes young children. SMARCB1, a core component of all SWI/SNF complexes, is biallelically inactivated in nearly all cases of RT, and studies have validated SMARCB1 as a potent and bona fide tumor suppressor. Inactivation of Smarcb1 through a reversibly inactivating conditional allele leads to cancer formation in mice with a median onset of only 11 weeks. In addition, mice with homozygous Smarcb1 deletions suffer
early embryonic lethality, and heterozygotes are predisposed to develop cancer. Interestingly, *SMARCB1*-deficient cancers are genomically stable, as human cancers with inactivated *SMARCB1* lack genomic amplifications or deletions. Furthermore, exome sequencing of RT shows that *SMARCB1* loss is the only recurrent mutation across many clinical samples. The mechanism by which SWI/SNF mutation leads to aggressive malignancy is currently not known. Given the quiet nature of the tumors harboring *SMARCB1* mutation, the study of *SMARCB1* mutation in cancer is a unique model for understanding both SWI/SNF biology and its role in cancer. The findings from studies of *SMARCB1* and RT will be extended to other SWI/SNF-mutant cancers; despite the rare nature of RT, the study of these cancers has the potential to impact the 20% of all cancers containing SWI/SNF mutations.

**SWI/SNF complexes and regulation of cancer-associated genes**

Many studies have elucidated pathways that are regulated by SWI/SNF complexes, and how disruption of these gene expression programs by subunit mutation can promote cancer. For example, SWI/SNF can bind to the retinoblastoma (RB) protein and facilitate repression of RB target genes. SWI/SNF also interacts with MYC, both as an activator and as a repressor. In part via disruption of RB function, inactivation of *SMARCB1* leads to downregulation of p16INK4a and E2F targets, indicating that SWI/SNF plays a role in cell cycle regulation and differentiation. Additionally, SWI/SNF complexes are required for specific regulation of interferon beta targets. It has also been shown that SWI/SNF complexes can bind to the promoters of roughly one-
third of all genes\textsuperscript{39,49} and the above represent only a few of the numerous pathways that have been shown to be dependent upon SWI/SNF.

An antagonistic relationship between SWI/SNF and the Polycomb repressive complex 2 (PRC2) has been described\textsuperscript{74,75}, providing an epigenetic mechanism to some of the tumor suppressive activity of SWI/SNF. SMARCB1 loss results in increased EZH2 levels as well as a concomitant increase in H3K27me3 at Polycomb target genes\textsuperscript{74}. This change in the chromatin landscape leads to repression of these genes, displaying an example of how SWI/SNF can affect gene expression at the epigenetic level. In addition, EZH2 inactivation blocks tumor formation in a mouse model of cancer driven by SMARCB1 loss\textsuperscript{74}, implicating EZH2 as an epigenetically-based therapeutic target for cancers harboring SWI/SNF mutation.

**Vulnerabilities of mutant SWI/SNF complexes in cancer**

One attractive hypothesis to account for many subunits of a single complex mutated in cancer is that all of the mutations are essentially equivalent and result in inactivation of SWI/SNF complexes. However, several findings conflict with such a possibility. First, the consequences of inactivation of genes encoding SWI/SNF subunits in mice are distinct. For example, while inactivation of Smarcb1 and Smarca4 both result in early embryonic lethality\textsuperscript{76–78}, knockout of Arid1a leads to the absence of mesoderm and arrest at E6.5\textsuperscript{28}, silencing of Smarcd3 results in heart developmental defects\textsuperscript{51}, and Smarca2-deficient mice are viable\textsuperscript{79}. Second, loss of different subunits of the complexes is associated with distinct cancer types. For example ARID1A is frequently mutated in ovarian cancer\textsuperscript{55,56}, SMARCA4 in lung cancer\textsuperscript{80,81}, PBRM1 in renal cancer\textsuperscript{82}, and
SMARCBI in rhabdoid tumors\textsuperscript{64}, with only modest overlap, suggesting specific consequences for mutation of different subunits. In addition, conditional Smarcb1 deletion in mice results in formation of rhabdoid-like tumors and lymphomas\textsuperscript{62} while Smarca4 haploinsufficiency leads to mammary tumors\textsuperscript{83}. Third, SWI/SNF complexes can assemble without SMARCBI\textsuperscript{84}, SMARCA4\textsuperscript{85,86} or ARID1A\textsuperscript{87}, indicating that residual complexes remain despite tumor suppressor subunit loss. Consequently, an alternate hypothesis was proposed: that loss of tumor suppressor subunits results in aberrant residual complexes that in turn actively drive oncogenesis\textsuperscript{88}. Furthermore, this aberrantly functioning residual complex could represent a new type of vulnerability in cancers containing SWI/SNF mutation.

Several recent publications support the idea of a residual complex as a vulnerability in SWI/SNF-mutant cancers. First, it has been shown that this residual complex is necessary for the proliferation of SMARCBI-mutant cancers\textsuperscript{88}. Specifically, the proliferation of SMARCBI-deficient human RT lines was blocked upon knockdown of SMARCA4, itself a tumor suppressor. In genetically engineered mouse models, inactivation of Smarca4 also blocked the \textit{in vivo} tumor formation otherwise caused by Smarcb1 loss. Data from additional papers has extended this concept from rare RT to more common cancers harboring mutations in other SWI/SNF subunits. For example, cancers mutant for ARID1A, the most frequently mutated subunit of the complex, were found to be specifically dependent upon ARID1B, a mutually exclusive member of the complex\textsuperscript{87}. This finding is described in detail in Chapter 2. Similarly, SMARCA4-mutant cancers were shown to be sensitive to loss of SMARCA2, the mutually exclusive catalytic subunit of SWI/SNF\textsuperscript{85,86,89}. Collectively, these recent findings strongly indicate
that residual complexes exist in a variety of SWI/SNF-mutant cancers and are essential for their growth (Figure 1-2)\textsuperscript{14}.

**Figure 1-2. Residual SWI/SNF complexes are a vulnerability in cancers containing SWI/SNF subunit mutation.** Mutation of a gene encoding a SWI/SNF complex subunit results in the formation of an aberrant residual complex, which is essential for the growth of the cancer. Targeting subunits of this residual complex is a newly identified therapeutic opportunity.
Role of aberrant DNA methylation in cancer

In addition to nucleosome remodeling and histone modification, DNA methylation is another type of epigenetic modification that affects transcription and gene expression. DNA is methylated when a methyl group is covalently added to the five-carbon position of cytosine bases in CpG dinucleotides. Most frequently, DNA methylation occurs on CpG islands and results in transcriptional repression, likely by blocking transcription factor binding to promoters. DNA methylation is catalyzed by two types of DNA methyltransferases (DNMTs). The first are de novo methyltransferases, such as DNMT3A and DNMT3B, which establish DNA methylation patterns during development. The second type are maintenance methyltransferases, such as DNMT1; these maintain the patterns of DNA methylation throughout many cell divisions. As a result, DNA methylation modifications are somatically heritable.

Abnormal CpG methylation patterns are associated with aberrant gene silencing in cancer, and both hypomethylation and hypermethylation occur in cancer cells. For example, normally expressed in neurons, the synuclein-γ gene SNCG is demethylated in breast and ovarian cancers and SNCG overexpression results in a more aggressive breast cancer phenotype. In contrast, tumor suppressor genes are often hypermethylated in cancer, resulting in their aberrant silencing. In tissues such as the kidney and pancreas, hypermethylation of genes is an early event in cancer progression, and an increase in methylation leads to a more cancerous state. Whether it be hypomethylation of oncogenes or hypermethylation of tumor suppressor genes, changes to DNA methylation are important contributors to cancer progression.
Previously, it has been shown that SWI/SNF complexes can affect gene expression at a specific locus by modulating DNA methylation. SWI/SNF evicts Polycomb proteins from the $INK4b$-$ARF$-$INK4a$ locus, decreasing DNA methylation and allowing for the transcriptional derepression of the genes $p16^{INK4a}$ and $p15^{INK4b}$\textsuperscript{75}. In addition, the catalytic subunits SMARCA4 and SMARCA2 have been implicated in the control of $CD44$ and $E$-cadherin gene expression through modulation of DNA methylation levels at the promoters of these genes\textsuperscript{97}. Taken together, these studies suggest that altering DNA methylation could be an important mechanism by which SWI/SNF complexes affect the epigenetic landscape in a cell.

References


Chapter 2

ARID1B is a specific vulnerability in ARID1A-mutant cancers
Abstract

Recent studies have revealed that *ARID1A* is frequently mutated across a wide variety of human cancers and also has *bona fide* tumor suppressor properties. Consequently, identification of vulnerabilities conferred by *ARID1A* mutation would have major relevance for human cancer. Here, using a broad screening approach, we identify ARID1B, a related but mutually exclusive homolog of ARID1A in the SWI/SNF chromatin remodeling complex, as the number one gene preferentially required for the survival of *ARID1A*-mutant cancer cell lines. We show that loss of ARID1B in ARID1A-deficient backgrounds destabilizes SWI/SNF and impairs proliferation. Intriguingly, we also find that *ARID1A* and *ARID1B* are frequently co-mutated in cancer, but that ARID1A-deficient cancers retain at least one *ARID1B* allele. These results suggest that loss of *ARID1A* and *ARID1B* alleles cooperatively promotes cancer formation but also results in a unique functional dependence. The results further identify ARID1B as a potential therapeutic target for *ARID1A*-mutant cancers.
Contributions:

The following individuals contributed to the work described in this chapter:

Katherine H. Walsh\textsuperscript{1,2,3,4*}, Xiaofeng Wang\textsuperscript{1,2,3*}, Boris G. Wilson\textsuperscript{1,2,3}, Francisca Vazquez\textsuperscript{5}, Jeffrey R. Haswell\textsuperscript{1,2,3}, Haley E. Manchester\textsuperscript{1,2,3}, Youngha Kim\textsuperscript{1,2,3}, Gregory V. Kryukov\textsuperscript{5}, Mahmoud Ghandi\textsuperscript{5}, Andrew J. Aguirre\textsuperscript{5,6,7}, Zainab Jagani\textsuperscript{8}, Zhong Wang\textsuperscript{9}, Levi A. Garraway\textsuperscript{6}, William C. Hahn\textsuperscript{6,7}, and Charles W. M. Roberts\textsuperscript{1,2,3,5,7}

\textsuperscript{1}Department of Pediatric Oncology, Dana-Farber Cancer Institute, Boston, MA 02215, USA
\textsuperscript{2}Division of Hematology/Oncology, Children's Hospital Boston MA, USA
\textsuperscript{3}Department of Pediatrics, Harvard Medical School, Boston MA, USA
\textsuperscript{4}Biological and Biomedical Sciences Program, Harvard Medical School, Boston MA, USA
\textsuperscript{5}Broad Institute of Harvard and Massachusetts Institute of Technology, Cambridge, Massachusetts, USA
\textsuperscript{6}Department of Medical Oncology, Dana-Farber Cancer Institute, Boston, MA 02215, USA
\textsuperscript{7}Center for Cancer Genome Discovery, Dana-Farber Cancer Institute, Boston, MA 02215, USA
\textsuperscript{8}Novartis Institutes for BioMedical Research, Cambridge, Massachusetts, USA.
\textsuperscript{9}Department of Cardiac Surgery, University of Michigan, Ann Arbor, MI 48109, USA

*These authors contributed equally to this work.

26
K. Walsh and X. Wang designed and performed experiments. J. Haswell and H. Manchester performed experiments. C. Roberts, K. Walsh, X. Wang, B. Wilson, J. Haswell, and H. Manchester analyzed and interpreted the data. K. Walsh, B. Wilson, F. Vazquez, and A. Aguirre performed analysis of Project Achilles data. G. Kryukov, M. Ghandi, and L. Garraway provided and analyzed sequencing data. Z. Wang provided Arid1a-conditional mice, intellectual contribution, and useful discussion. Y. Kim established Arid1a-conditional mouse strain. W. Hahn directs the Achilles Project, provided reagents, helped interpret Achilles data, and edited the manuscript. C. Roberts, K. Walsh, and X. Wang wrote the manuscript.

Note: This entire chapter, except for the section on in vivo xenograft experiments, was published in the article: Helming KC*, Wang X*, et al. ARID1B is a specific vulnerability in ARID1A-mutant cancers. Nat. Med. 20, 251-254 (2014).
Results

*ARID1B is specifically essential for ARID1A-mutant cells*

To search for specific dependencies created by *ARID1A* mutation, we utilized data from Project Achilles, a large-scale project focused on identifying essential genes in a wide panel of cancer cell lines using genome-scale loss-of-function genetics\(^1\). We compared 18 *ARID1A*-mutant and 147 cell lines wildtype for *ARID1A* (hereafter referred to as wildtype). Of 9,050 genes interrogated, *ARID1B* scored as the top candidate preferentially required for the growth of *ARID1A*-mutant cancer cell lines (p=7.366x10\(^{-6}\), FDR<0.001) (Figure 2-1A, C). Vulnerability to ARID1B depletion was even more pronounced in the large subset of cell lines that contained inactivating *ARID1A* mutations (rather than missense mutations) (Figure 2-1B), suggesting that ARID1B is specifically essential for cell lines lacking ARID1A. To further evaluate this finding, we examined effects of the individual ARID1B shRNAs. Three of the four ARID1B shRNAs passed the Achilles quality control metrics\(^2\). These scored #1 (p=1.211x10\(^{-6}\), FDR<0.001), #4 (p=1.211x10\(^{-6}\), FDR<0.001), and #11 (p=1.816x10\(^{-5}\), FDR=0.090) of the 54,020 shRNAs in the screen. We confirmed ARID1B as a vulnerability by knocking it down in two cell lines that contained *ARID1A*-inactivating mutations, OVISE and TOV21G, and two *ARID1A* wildtype lines, ES-2 and 293T (Figure 2-2A). Proliferation (Figure 2-2B) and colony formation (Figure 2-2C) were impaired in *ARID1A*-mutant cells but not in wildtype cells.
Figure 2-1. Genome-wide shRNA screen identifies ARID1B as a specific vulnerability in ARID1A-mutant cancer cell lines. A. Rank list of vulnerabilities identified by screen of Achilles platform cell lines. ARID1B is the #1 gene preferentially essential for the growth of ARID1A-mutant cancer cell lines as compared to wildtype cancer cell lines. B. Effects of ARID1B shRNAs across cell lines in the Achilles screen. Negative values indicate growth inhibition while positive values reflect growth enhancement. In the top panel, cell lines with any identified ARID1A mutation compared to the reference genome are indicated in red. In the bottom panel, only those cell lines with clear inactivating mutations in ARID1A are shown in red. C. Position of SWI/SNF subunits and top 5 genes in the class comparison are indicated on curve.
Figure 2-2. ARID1B loss specifically impairs the proliferation of *ARID1A*-mutant cancer cells but not wildtype cells. A. Immunoblots showing the results of two independent shRNAs targeting ARID1B in the ES-2, 293T, OVISE, and TOV21G cell lines. B. Proliferation of wildtype (ES-2 and 293T) and *ARID1A*-mutant (OVISE and TOV21G) cell lines in response to two independent ARID1B shRNAs. * p < 0.0002; ** p < 4x10^{-8}; NS, p > 0.05; # p < 0.05. Data are expressed as mean ± S.D. C. Colony formation in response to ARID1B knockdown in wildtype and *ARID1A*-mutant cell lines.
Of the cell lines in the Achilles screen, 18 contained ARID1A inactivating mutations, four contained ARID1A missense mutations, and eight had ARID1A 3’ UTR mutations. Of note, it is not known whether the missense mutations impair ARID1A function and act as drivers, or whether due to the high rate of mutations in some cancers, or acquired in cell culture, these mutations do not affect ARID1A function and are passenger mutations. The top panel of Figure 2-1B shows the effect of ARID1B knockdown in all 30 Achilles cell lines carrying any ARID1A mutation (inactivating, missense, or 3’ UTR). The bottom panel shows the effect upon only those 18 cell lines that carry inactivating mutations. As can be seen by comparing the two, the lines with missense mutation appear unaffected by ARID1B knockdown. It is unknown whether this lack of dependence occurs because only inactivating ARID1A mutations result in dependence upon ARID1B or because the missense mutations are passenger mutations and would not be expected to result in dependence upon ARID1B.

ARID1B loss destabilizes SWI/SNF complexes in ARID1A-mutant cells

ARID1A and ARID1B are 60% identical, have been reported to have opposing functions in cell cycle arrest, and are mutually exclusive since individual SWI/SNF chromatin remodeling complexes can contain either ARID1A or ARID1B, but not both. To investigate the relationship between ARID1A and ARID1B in cancer, we asked whether an ARID1B-containing SWI/SNF complex was present in ARID1A-mutant cells. Immunoprecipitation of the SMARCC1 (BAF155) core subunit of the SWI/SNF complex resulted in co-precipitation of ARID1B and other SWI/SNF subunits in both wildtype and ARID1A-mutant cells, indicating that intact ARID1B-containing complexes are present (Figure 2-3A, B, C) in both wildtype and ARID1A-mutant cells.
Figure 2-3. ARID1B is required for the maintenance of an intact SWI/SNF complex in ARID1A-mutant cancer cell lines. A-C. Immunoprecipitation of the SWI/SNF complex by SMARCC1 from the nuclear extract of ES-2 (A), OVISE (B) and TOV21G (C) cells upon treatment with control shRNA or two independent ARID1B shRNAs. D. RT-qPCR analysis of the expression of indicated SWI/SNF subunits in ES-2, OVISE, and TOV21G cells with either control shRNA or ARID1B shRNAs.
We next sought to determine whether ARID1B loss affects the composition of the SWI/SNF complex in ARID1A-mutant cancer cells. Knockdown of ARID1B in wildtype cell lines had no effect on the expression of other SWI/SNF complex subunits or upon their incorporation into the complex (Figure 2-3A). However, depletion of ARID1B in ARID1A-mutant cells resulted in dissociation of the core catalytic ATPase subunit SMARCA4 (BRG1) and reduced incorporation of several other subunits (Figure 2-3B, C). Protein levels of core subunits such as SMARCA4, SMARCC2, and SMARCB1 were also decreased, particularly in the TOV21G line (Figure 2-3C), while the mRNA levels were largely unaffected (Figure 2-3D), suggesting post-translational loss of these proteins.

To further investigate how ARID1B loss affects the assembly of the SWI/SNF complex, we performed a sucrose sedimentation assay on cells treated with either control shRNA or ARID1B shRNA. Consistent with the co-immunoprecipitation results, an intact 2 MDa SWI/SNF complex is observed in ARID1A-mutant cells treated with control shRNA (Figure 2-4) and in wildtype cells treated with either control or ARID1B shRNA (Figure 2-5). In contrast, knockdown of ARID1B in ARID1A-mutant cells eliminated the intact SWI/SNF complex (Figure 2-4), indicating that in human ARID1A-mutant cancer cell lines, the presence of ARID1B is essential for formation or stabilization of an intact SWI/SNF complex. Despite ARID1A/ARID1B and PBRM1 having been reported to exist in mutually exclusive versions of the SWI/SNF complex, our findings are consistent with a more recent publication, which found that these subunits can co-exist, as PBRM1 association with smaller complexes was substantially affected by the combined loss of ARID1A/B (Figure 2-4).
Figure 2-4. SWI/SNF complex is destabilized upon ARID1B loss in *ARID1A*-mutant cells. Sucrose sedimentation (20-50%) assay of SWI/SNF complexes in *ARID1A*-mutant cells treated with either control or ARID1B shRNA in OVISE (A) or TOV21G (B) cells.
Figure 2-5. ARID1B loss does not affect the SWI/SNF complex in ARID1A WT cells.

Sucrose sedimentation assay of SWI/SNF complexes in ES-2 cells treated with either control shRNA or ARID1B shRNA.
As the SWI/SNF complex binds up to one-third of all genes\(^7\) and several members of the SWI/SNF complex are essential in mouse development\(^8\)–\(^10\) and for survival of many cell lineages\(^11\)–\(^12\), loss of an intact SWI/SNF complex would be predicted to be incompatible with cell viability.

In order to further validate the identification of ARID1B as a vulnerability in ARID1A-mutant human cancer, we sought to investigate whether inactivation of Arid1a creates a dependence upon ARID1B using primary mouse embryonic fibroblasts (MEFs) conditional for Arid1a\(^13\). Deletion of Arid1a or knockdown of ARID1B individually had only moderate effects on proliferation, while combined loss led to substantial impairment (Figure 2-6A). We similarly observed that loss of Arid1a or ARID1B alone had only modest effects on the composition of the complex, while loss of both led to dissociation and degradation of SMARCA4 and substantial reductions in stability and incorporation of many other SWI/SNF subunits (Figure 2-6B, D). Again, the reduced protein levels were not due to changes in transcription (Figure 2-6C). Finally, sucrose sedimentation assay showed that loss of Arid1a and ARID1B in MEFs eliminated the intact SWI/SNF complex (Figure 2-6D).
Figure 2-6. *Arid1a* loss creates a dependency on ARID1B-containing SWI/SNF complexes in primary cells. A. MTT proliferation assay of control MEFs, *Arid1a* knockout (KO) MEFs, ARID1B knockdown (KD) MEFs, or combined *Arid1a* KO and ARID1B KD MEFs. B. Immunoprecipitation of the SWI/SNF complex by SMARCC1 from the nuclear extract of MEFs with indicated treatment. C. mRNA levels of the SWI/SNF complex subunits upon individual loss of *Arid1a*, ARID1B or both. D. Sucrose sedimentation (20-50%) assay of the SWI/SNF complex from the nuclear extract of MEFs with indicated treatment: control shRNA treated MEFs (top half) or *Arid1a* KO and ARID1B KD MEFs (bottom half).
ARID1A and ARID1B mutations co-occur in cell lines and primary cancers

Collectively, these findings demonstrate a synthetic lethal relationship between this mutually exclusive pair of SWI/SNF subunits. Notably, however, ARID1B has also been reported mutated in human cancers\textsuperscript{3,14}, and has been found to be mutant in some of the same types of cancer as ARID1A, such as neuroblastoma\textsuperscript{14}. Since we found ARID1B knockdown to impair the growth of ARID1A-mutant cell lines, we initially hypothesized that mutations in ARID1A and ARID1B would be mutually exclusive. Surprisingly, we found that ARID1A and ARID1B mutations co-occur in both cancer cell lines and primary tumors. Using data from cell line sequencing\textsuperscript{15,16} we found that 38\% of 34 ARID1A-mutant lines also contained ARID1B-inactivating mutations\textsuperscript{16} (p<1x10\textsuperscript{-6}). Notably, all lines retained at least one allele of either ARID1A or ARID1B, suggesting that retention of at least one ARID1 allele may be essential for survival. This finding also held true in primary cancer samples. We found that of the 297 ARID1A-mutant primary cancer samples cataloged in the cBio Portal for Cancer Genomics\textsuperscript{17,18}, 30 (10.1\%) also contained ARID1B mutations (p=1.07x10\textsuperscript{-7}), significantly higher than the 3\% rate in ARID1A-wildtype tumors.

The co-occurrence of ARID1A and ARID1B mutations raises the possibility that the synthetic lethality relationship could be caused simply by the high frequency of ARID1B mutations in ARID1A-mutant cancer cell lines. To evaluate this possibility, we removed all ARID1B-mutant cell lines and conducted a revised class comparison in which four ARID1B-wildtype, ARID1A-mutant cell lines were compared to 49 cell lines wildtype for both ARID1A and ARID1B. ARID1B still scored number four out of 9,000+
genes (p=7.154x10^{-4}), indicating that the synthetic lethality between ARID1A and ARID1B is a result of ARID1A mutation and not co-occurring ARID1B mutations.

**In vivo xenograft experiments**

In order to further validate and extend the finding of ARID1A and ARID1B synthetic lethality to an animal model, we conducted a series of *in vivo* xenograft experiments. In concordance with the Project Achilles and *in vitro* data, we hypothesized that knockdown of ARID1B would block the growth of ARID1A-mutant tumors *in vivo*. Despite performing many iterations of this experiment, the technical challenges prevented strong conclusions from being drawn regarding ARID1A and ARID1B synthetic lethality *in vivo*.

The first experimental set up utilized constitutive ARID1B shRNAs and non-silencing shRNAs against LacZ or GFP as a control. ARID1A-mutant TOV21G cells were treated with shRNAs before injection into nude mice. In three independent experiments, cells treated with ARID1B shRNA surprisingly formed larger tumors than the cells treated with control shRNA (Figure 2-7A). It was observed, however, that the control shRNAs against LacZ and GFP both displayed off target effects, as the cells treated with these shRNAs exhibited growth impairment in cell culture (Figure 2-7B). In addition, while the ARID1B shRNAs moderately reduced ARID1B protein expression, the tumors retained partial expression of ARID1B (Figure 2-7C, D). We hypothesized that this partial retention of ARID1B expression, combined with off target effects from the control shRNA, accounted for the observed discrepancy between the *in vitro* and *in vivo* data.
Figure 2-7. *In vivo* xenograft experiment with constitutive ARID1B knockdown.

A. Histogram displaying average tumor volume at the end point of the experiment.

B. Colony forming assay of TOV21G cells treated with shRNAs grown in parallel with xenograft studies. C. Western blot of nuclear extract from tumors.

D. Immunohistochemistry of tumor with ARID1B antibody (Abcam).
In attempt to circumvent the technical issues described above, we utilized a doxycycline-inducible shRNA system to knock down ARID1B in a much more precise and controlled manner. This system is internally controlled, as animals treated with doxycycline can be directly compared to animals not receiving doxycycline treatment, effectively removing the need for a potentially confounding non-silencing control. TOV21G cell lines were established with an inducible ARID1B shRNA and it was confirmed that treatment of these cells in vitro resulted in effective suppression of ARID1B. These cells were then injected subcutaneously into nude mice. Once the tumors formed to an approximate volume of 100mm³, the animals were randomized into doxycycline or control treatment groups. The experiment ended when the tumors reached an approximate volume of 1000mm³, and the tumors were analyzed for ARID1B expression.

The volume and mass of the tumors were approximately identical when comparing the doxycycline-treated animals to the control animals (Figure 2-8A, B). Immunohistochemistry and Western blot analysis in nuclear extract from the tumors revealed no clear difference in ARID1B expression between the induced and uninduced samples and also showed a high degree of variability (Figure 2-8C, D), likely a reflection of an impure tumor sample with contamination from surrounding tissue. RT-qPCR shows that RNA levels of ARID1B are slightly decreased in the doxycycline-treated samples as compared to control samples (Figure 2-8E). The same results were observed in tumors formed from OVISE cells, an additional and distinct ARID1A-mutant human ovarian cancer cell line.
Figure 2-8. *In vivo* xenograft experiment with inducible ARID1B knockdown.

A. Approximate tumor volume as measured by calipers on day 32 post-injection (the day the mice were sacrificed). Graph displays median value and interquartile range. B. Mass of tumor following dissection. Graph displays median value and interquartile range. C. Immunohistochemistry of tumors with ARID1B antibody (Abcam). D. Western blot of nuclear extract from tumors. E. RT-qPCR from tumors using two primer pairs against ARID1B.
The discordant results between the *in vitro* data and the *in vivo* experiments could be explained by several possibilities. First, it is likely that the inducible ARID1B shRNA does not sustain sufficient protein knockdown long enough to cause a growth difference *in vivo*. An alternative possibility is that there is a strong selective pressure *in vivo* for the cells to maintain ARID1B expression and that this retention of ARID1B allows the shRNA-treated cells to grow faster than other cells. It is also possible that subcutaneous xenograft injection is not an effective system for testing this hypothesis. The intrinsic variability of this experimental set up makes it difficult to draw strong conclusions from the data. For example, despite careful quality control measures, the tumor size consistently ranges widely in both experimental conditions. It will be important to validate this relationship *in vivo* with improved reagents, such as cells with CRISPR knockout of *ARID1B* as opposed to shRNA-mediated knockdown.

**Conclusion**

In this report, we show that inactivating mutations in *ARID1A*, frequent across a wide variety of cancers, create a dependency upon ARID1B (Figure 2-9). It is notable that the number one vulnerability in *ARID1A*-mutant cell lines is another member of the SWI/SNF complex. We previously showed that cancer formation in the absence of the SWI/SNF subunit SMARCB1 does not result from SWI/SNF inactivation but rather that oncogenesis was dependent upon the activity of the residual SWI/SNF complex19. At that time we speculated that, much like the concept of oncogene addiction, targeting the aberrant residual SWI/SNF complex might theoretically be an effective therapeutic approach for *SMARCB1*-mutant cancers. Our present study, which surveyed 9,050 genes,
reveals the role of the residual complex in the growth of *ARID1A*-deficient cancers and also identifies a specific subunit as a relative vulnerability. This principle may have broad applicability to SWI/SNF-mutant cancers as Oike et al recently showed that SMARCA2, a paralog of SMARCA4, was essential in *SMARCA4*-mutant cancers\(^20\).

**Figure 2-9. Model: Inactivating mutations in *ARID1A* promote oncogenic transformation but also create specific dependency on ARID1B.** Inhibition of ARID1B in *ARID1A*-mutant cells destabilizes the SWI/SNF complex and results in impaired cell growth.

Together, our results may suggest that partial loss of ARID1 function via mutation of *ARID1A* alleles or, less frequently, *ARID1B* alleles can drive cancer growth but at the same time create a specific vulnerability compared to non-mutant cells. This finding suggests ARID1B as a potential therapeutic target for cancers that contain inactivating *ARID1A* mutations. Recent examples have demonstrated the feasibility and efficacy of targeting chromatin regulators such as BRD4\(^{21,22}\) as well as other non-enzymatic proteins such as BCL-2\(^{23}\) and molecules previously found difficult to target such as RAS\(^{24}\). ARID1B could potentially be targetable through its E3 ubiquitin ligase
interaction\textsuperscript{25}. Additionally, novel approaches using small stabilized peptides have recently been shown capable of displacing the EZH2 subunit from its Polycomb PRC2 chromatin modifying complex\textsuperscript{26}. Analogous approaches may now be considered for targeting ARID1B.

**Methods:**

**Achilles Analysis**

To find genes that are preferentially essential in mutant cell lines, we used the GenePattern module PARIS (http://www.broadinstitute.org/cancer/software/genepattern) using the default parameters except quality, which was changed to final. The gene-level Achilles dataset v2.4 was used as a data file (www.broadinstitute.org/achilles) (file name: Achilles_QC_v2.4_rnai.Gs.gct). The classifier files were generated using the gene mutation status from the Cancer Cell Line Encyclopedia (www.broadinstitute.org/ccle). Cell lines without hybrid capture sequencing data were removed from the analysis. The mutational status of *ARID1A* was annotated for 165 of the 216 cell lines in the Achilles dataset, and as a result, these 165 cell lines were used in the class comparisons.

**Cell Culture**

TOV21G (CRL 11730), ES-2 (CRL-1987), and 293T (CRL-3216) cell lines were purchased from ATCC. OVISE cells were obtained from William Hahn’s laboratory. Mouse Embryonic Fibroblasts (MEFs) were generated as described previously\textsuperscript{27}. Cells were transduced with shRNAs and selected with puromycin for 48-72 hours before seeding for MTT or colony formation assays. MTT assays were conducted with a Cell Proliferation Kit (Roche). Colony formation assays were conducted by staining cells for
20 minutes with crystal violet staining solution (0.05% Crystal Violet, 1% Formaldehyde, 1% PBS, 1% methanol).

**shRNA-mediated knockdown of ARID1B**

ARID1B shRNAs were obtained from the RNA interference (RNAi) screening facility at the Dana-Farber Cancer Institute and were lentivirally transduced into OVISE, TOV21G, ES-2, and 293T cells. ARID1B and non-silencing control shRNAs are in the pLKO.1 lentiviral expression vector backbone. Target sequence for ARID1B shRNA#1 (361) is GCCGAATTACAAACGCCATAT and the target sequence for ARID1B shRNA#2 (363) is GCAGTATATTCAGTACCTGTT.

**Density Sedimentation Analysis**

Nuclear extract (1 mg) was diluted in 300 µl of 0% sucrose RIPA buffer and carefully overlaid onto a 12 ml 20%–50% sucrose (in RIPA buffer) gradient prepared in a 14ml 14×95 mm polyallomer centrifuge tube (Cat. #331374, Beckman Coulter). Tubes were placed in an SW-40 Ti swing bucket rotor and centrifuged at 4°C for 16 hours at 40,000 rpm. Fractions (0.5 ml) were collected and used in gel electrophoresis and subsequent Western blotting analyses.

**Immunoblots and co-immunoprecipitation experiments**

Whole cell extracts for immunoblotting were prepared by incubating cells on ice in 1% NP-40 lysis buffer (50 mM Tris-HCL pH 7.4, 5 mM EDTA, 12% Glycerol, 50 mM NaCl, 1% NP-40) plus protease inhibitors (Complete, Mini, EDTA-free. Roche: 11836170001) for 30 minutes. Supernatants were collected following a brief spin (10 min.) at 17900 r.c.f. to separate cellular debris in a 4°C centrifuge. Protein concentrations were determined using the Bradford reagent (Biorad). SDS-
polyacrylamide gel electrophoresis was used to separate proteins, which were subsequently transferred to PVDF membranes (Millipore). ARID1B antibody (Abcam: ab54761) was used to detect efficient knockdown.

Nuclear extracts for immunoprecipitation (IP) were prepared using the NE-PER Nuclear and Cytoplasmic Extraction Kit (Thermo Scientific: 78833). Nuclear extracts were diluted with RIPA buffer (1 mg/ml, with protease inhibitors and DTT). Each IP was incubated with indicated antibodies overnight at 4° C. Protein G Dynabeads (Life Technologies: 10004D) were added and incubated at 4° C for 3 hrs. Beads were then washed three times with RIPA buffer and resuspended in reducing SDS gel loading buffer. Antibodies used in the immunoprecipitation and immunoblots are:

- SMARCC1/BAF155 (Santa Cruz: 9746); ARID1A (Bethyl Laboratories: A301-041A);
- PBRM1 (Bethyl Laboratories: A301-591A); SMARCA4 (Santa Cruz: sc17796);
- SMARCC2/BAF170 (Bethyl Laboratories: A301-039A); SMARCD1/BAF60A (Bethyl Laboratories: A301-595A); SMARCE1/BAF57 (Bethyl Laboratories: A300-810A);
- ACTL6A/BAF53A (Bethyl Laboratories: A301-391A); ACTIN (Cell Signaling Technology: 5125).

**RNA purification and RT-qPCR**

Total RNA was extracted using Trizol reagent (Invitrogen) following the manufacturer’s instructions. 2 µg of total RNA was reverse-transcribed into first-strand cDNA using oligo(dT)$_{20}$ primers and the SuperScript II or SuperScript III Reverse Transcriptase (Invitrogen). RT-qPCR was performed on the ViiA 7 Real-Time PCR System (Life Technologies) using SYBR Select Master Mix (Life Technologies).
Reactions were performed in triplicate, and gene expression was normalized to GAPDH.

Error bars represent SD of mean expression.

**Cell Line Sequencing**

Cell lines were sequenced as previously described\(^\text{15}\). Cell line sequencing data and the data from the Cancer Cell Line Encyclopedia\(^\text{16}\) were used to identify cell lines with co-occurring mutations of \(ARID1A\) and \(ARID1B\).

**Statistical Significance of Mutation Overlap**

To evaluate the statistical significance of the overlap of \(ARID1A\) and \(ARID1B\) mutations, the probability of observing at least \(n_{12}\) cell lines with both mutations was estimated under the null hypothesis that these two mutations are independent. For that, given \(n_1\) cell lines with \(ARID1A\) and \(n_2\) cell lines with \(ARID1B\) mutations and \(n_{12}\) cell lines with both mutations the following simulation was run: \(n_1\) cell lines were randomly picked with the probability for each cell line being selected set relative to its mutation rate and assigned mutation ‘A’ to these cell lines. Next, \(n_2\) cell lines were similarly selected and assigned mutation ‘B’ to those cell lines and then the number of cell lines with both mutations ‘A’ and ‘B’ was counted. This process was repeated many times to estimate the probability of observing \(n_{12}\) cell lines or more with both mutations.

For primary cancer samples, a contingency table was formed consisting of the counts for all the four possibilities of \(ARID1A\) or \(ARID1B\) mutation status. The Fisher’s exact test was used to calculate the statistical significance of the overlap of \(ARID1A\) and \(ARID1B\) mutations.
Xenograft experiments

Nude mice were ordered from Charles River Labs (strain 088). TOV21G and OVISE cells (described above) were lentivirally infected with constitutive or inducible ARID1B shRNA #1 as described above. Following selection, 2.4 or 5 million cells resuspended in PBS + 5% FBS were mixed with matrigel (BD Biosciences 356231) and injected subcutaneously into the flank of nude mice. For inducible experiments, mice were randomized to doxycycline-containing food (625ppm) upon formation of tumors measuring approximately 100mm$^3$. Mice were monitored three times per week, tumor size was measured using calipers, and the equation $[(\text{width}^2 \times \text{length})/2]$ was used to approximate volume. The animals were sacrificed when the tumor reached a volume of approximately 1000mm$^3$ or the animal appeared moribund, whichever came first. Following euthanasia, the tumors were harvested and dissected into three parts. The first part was fixed in 10% neutral buffered formalin and was sent to the Specialized Histopathology Core at Brigham and Women’s hospital for immunohistochemistry of ARID1B (Abcam antibody 54761). The second piece was disassociated in Trizol (Life Technologies) for RNA analysis. The third and final piece was homogenized using a glass tissue homogenizer and the nuclear extract was obtained for protein analysis as described above.

References


Chapter 3

Evaluation of the redundancy between ARID1A and ARID1B
Abstract

The findings of synthetic lethality between ARID1A and ARID1B, combined with the high frequency of co-mutations, present an unexpected and complicated interplay between these two highly homologous members of the SWI/SNF complex. In order to further characterize the roles of ARID1A and ARID1B, I evaluated their functions in cell proliferation, SWI/SNF complex composition, and gene expression. I hypothesized that ARID1A and ARID1B would serve unique functions but that they also would retain redundant roles. To reconcile the seemingly opposite observations of synthetic lethality and co-mutation, I also hypothesized that a dosage relationship exists between ARID1A and ARID1B such that partial loss of both leads to a growth advantage while complete loss is incompatible with cell viability. I tested these hypotheses via genetic knockout of \textit{ARID1A} and \textit{ARID1B} alleles coupled with shRNA-mediated protein knockdown of ARID1A and ARID1B. These results have revealed a fascinating but complicated relationship between these two proteins, suggesting that cell lineage and mutational background may play key roles in determining the relationship between ARID1A and ARID1B.
Contributions:

The following individuals contributed to the work described in this chapter:

Katherine H. Walsh\textsuperscript{1,2,3,4}, Haley E. Manchester\textsuperscript{1,2,3}, Weishan Wang\textsuperscript{1,2,3}, Burak Alver\textsuperscript{5}

\textsuperscript{1}Department of Pediatric Oncology, Dana-Farber Cancer Institute, Boston MA 02215, USA
\textsuperscript{2}Division of Hematology/Oncology, Children's Hospital Boston MA, USA
\textsuperscript{3}Department of Pediatrics, Harvard Medical School, Boston MA, USA
\textsuperscript{4}Biological and Biomedical Sciences Program, Harvard Medical School, Boston MA, USA
\textsuperscript{5}Center for Biomedical Informatics, Harvard Medical School, Boston MA, USA

K. Walsh designed and performed all experiments. H. Manchester and W. Wang provided technical assistance. B. Alver analyzed gene expression data.
Introduction

The results described in Chapter 2 present a complicated relationship between these two homologues of SWI/SNF complexes. It is unexpected to find two synthetic lethal genes also co-mutated, as it would be predicted that inactivating mutations, and therefore loss of both genes, would be detrimental to the cells. An alternative explanation to account for synthetic lethality and co-mutation is that cells gain a growth advantage as they lose alleles of *ARID1A* and *ARID1B*, but that complete loss of both genes is lethal. In this scenario, co-mutation of *ARID1A* and *ARID1B* would be evolutionarily beneficial to the cells, but complete loss of both proteins would be synthetic lethal.

Additionally, the precise role of each protein, and their degree of functional redundancy, is not fully understood. As detailed in Chapter 1, ARID1A and ARID1B have been reported to have unique roles, such as in cell cycle control\(^1\), but whether their functions are entirely distinct remains to be determined. It is possible that ARID1A and ARID1B can partially substitute for each other in certain contexts or developmental states, but that they also retain unique and specific functions. To test this hypothesis, several experiments have been conducted to explore the roles of these proteins on cell proliferation, SWI/SNF complex composition, and gene expression. These studies illuminate a fascinating, though complex, dynamic between ARID1A and ARID1B, which has both interesting biological relevance and important therapeutic implications.
Results

CRISPR-mediated deletion of ARID1B in ARID1A isogenic cell lines

To study the dosage relationship between ARID1A and ARID1B, I utilized HCT116 ARID1A isogenic cell lines purchased from Horizon Discovery. Isolated from a human large intestine carcinoma, this ARID1A wildtype (WT) cell line has been engineered to have a premature stop codon knocked in at position Q456, located at the start of exon 3 in the ARID1A gene. This stop codon caused either heterozygous (het) or homozygous (hom) ARID1A deletion. In order to test the hypothesis that the cell lines would acquire a growth advantage as they lost alleles of ARID1 (either A or B) until they reached a level of ARID1 alleles that would not be tolerable, I sought to use CRISPR$^{2,3}$ to generate ARID1B het and hom deletions in the ARID1A WT, het, and hom cell lines. The goal was to generate cell lines with every possible combination of ARID1A and ARID1B alleles and to measure the combinatorial effects on cell growth and SWI/SNF complex composition.

In order to delete ARID1B, guide sequences were designed that recognized a sequence preceding the promoter of ARID1B and an intronic sequence following exon 1. Having been guided to these recognition sequences, Cas9 generated double stranded breaks that then triggered the non homologous end joining (NHEJ) machinery to repair the breaks, thereby removing the DNA in between the two cut sites$^4$ (Figure 3-1). I transfected the CRISPR/Cas9 plasmids with the ARID1B guide sequences into the ARID1A isogenic cell lines, sorted the cells, plated single-cell clones, and screened these clones for deletion by PCR. This strategy effectively deleted the promoter and exon 1 of ARID1B, resulting in complete loss of ARID1B gene expression.
Figure 3-1. Schematic diagram of CRISPR deletion strategy. Guide sequences direct Cas9 to initiate two double stranded breaks, which when repaired with NHEJ, result in a deletion of the intervening sequence. Distance between cut 1 and cut 2 is approximately 10 kilobases.

Clones of all genotypes, except a double hom clone, were isolated (Table 3-1), and ARID1B protein (Figure 3-2A) and RNA levels (Figure 3-2B) were confirmed to be depleted. The growth phenotype of each cell line is summarized in Table 3-1 and displayed in Figure 3-2C. Interestingly, in the ARID1A hom, ARID1B het clone, ARID1B protein and RNA levels are not detected, possibly identifying a clone that has lost protein expression of both ARID1A and ARID1B. While supportive of the synthetic lethal relationship, a growth advantage upon partial loss of ARID1A and ARID1B alleles is not observed, suggesting that the dosage hypothesis may be incorrect, or that the relationship between ARID1A and ARID1B is cell context dependent and is different in this cell line.
Table 3-1. Summary of growth of CRISPR clones.

<table>
<thead>
<tr>
<th>Genotype</th>
<th>Growth relative to control line</th>
<th>% of clones screened</th>
</tr>
</thead>
<tbody>
<tr>
<td>ARID1A WT, ARID1B HET</td>
<td>Normal growth</td>
<td>27.6</td>
</tr>
<tr>
<td>ARID1A WT, ARID1B HOM</td>
<td>Normal growth</td>
<td>2.1</td>
</tr>
<tr>
<td>ARID1A HET, ARID1B HET</td>
<td>Normal growth</td>
<td>11.9</td>
</tr>
<tr>
<td>ARID1A HET, ARID1B HOM</td>
<td>Slightly impaired growth</td>
<td>4.7</td>
</tr>
<tr>
<td>ARID1A HOM, ARID1B HET</td>
<td>Significantly impaired growth</td>
<td>1.4</td>
</tr>
<tr>
<td>ARID1A HOM, ARID1B HOM</td>
<td>Clone not generated</td>
<td>0</td>
</tr>
</tbody>
</table>
Figure 3-2. *ARID1B* CRISPR in *ARID1A* isogenic cell lines. A. Western blot confirming protein loss in *ARID1B* homozygous clones. B. RT-qPCR displaying *ARID1B* RNA levels following CRISPR deletion. C. Colony forming assay showing the effect of *ARID1B* deletion on cell growth.
Figure 3-2 (Continued)

A.  

<table>
<thead>
<tr>
<th></th>
<th>ARID1A WT</th>
<th>ARID1A HET</th>
<th>ARID1A HOM</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1B</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WT</td>
<td>A1B WT</td>
<td>A1B HET</td>
<td>A1B HOM</td>
</tr>
<tr>
<td></td>
<td>A2  A3 A5 A6 A4 D9 D12 D9 D7 2D9 4E3 4E7 4B4</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

250-  

ARID1A

250-  

ARID1B

ACTIN

B.  

![Graph showing ARID1A WT, ARID1A HET, ARID1A HOM bars with error bars for A1B WT, A1B HET, A1B HOM conditions with two markers ARID1B_1 and ARID1B_2.]

ARID1A WT  ARID1A HET  ARID1A HOM

C.  

Control (no deletion)  Heterozygous ARID1B deletion  Homozygous ARID1B deletion
Immunoprecipitation of SMARCC1, a core SWI/SNF subunit, was conducted to examine the composition of SWI/SNF complexes upon loss of ARID1A and ARID1B alleles. Overall, the SWI/SNF complex remained intact following deletion of ARID1A and ARID1B, except in the ARID1A hom, ARID1B het clone where the incorporation of the catalytic subunit SMARCA4 and variant subunit ACTL6A are decreased (Figure 3-3). It is important to note that ARID1A expression appears to be regained in the ARID1A hom cells. It is possible that these clones may have adapted to accommodate ARID1B loss, resulting in concomitant changes that alter the expression of other genes. For example, ARID1A protein levels may be regained in the ARID1A hom cells, perhaps by circumventing the stop codon, in order to cope with ARID1B loss. An additional possibility is that certain isoforms of ARID1A may be retained or upregulated upon ARID1B CRISPR deletion. Taken together, these results support the previous data showing that complete loss of both ARID1A and ARID1B destabilizes SWI/SNF complexes, but provides new information that partial loss of either does not affect the complex in this cell line.
Figure 3-3. Composition of SWI/SNF complexes in *ARID1A* isogenic cell lines with *ARID1B* CRISPR deletion. Immunoprecipitation of SMARCC1 from the nuclear extract of HCT116 *ARID1A* isogenic cells with *ARID1B* CRISPR deletion.
shRNA-mediated knockdown of ARID1A and ARID1B

To complement the CRISPR studies and further dissect the intricate relationship between these two proteins, I utilized shRNAs to modulate the expression of ARID1A and ARID1B at the protein level. First, I knocked down ARID1B in the HCT116 ARID1A isogenic cell lines (Figure 3-4A) and measured cell proliferation and colony forming ability (Figure 3-4B, C). I hypothesized that ARID1B loss would have little or no effect on the ARID1A wildtype cell line and that ARID1B loss would impair the growth of the ARID1A hom cell line. If the dosage hypothesis is correct, I also expected that the ARID1A het cell line would grow better than the ARID1A wildtype cell line upon ARID1B knockdown. The results from this experiment, however, indicate that the HCT116 ARID1A wildtype cells are sensitive to ARID1B loss and the ARID1A het and hom cells lines are markedly affected by ARID1B loss (Figure 3-4B, C). Interestingly, ARID1A protein expression increases in the ARID1A heterozygous cell line upon ARID1B knockdown, suggesting that these cells are trying to compensate for ARID1B loss by upregulating ARID1A (Figure 3-4A).
Figure 3-4. ARID1B knockdown in HCT116 ARID1A isogenic cell lines. A. Western blot of HCT116 isogenic cell lines infected with inducible ARID1B shRNA. B. Colony forming and C. MTT assays of isogenic cell lines following ARID1B knockdown.
Consistent with the results from the CRISPR experiments, these data do not show an acquired growth advantage upon partial loss of *ARID1A* and *ARID1B* alleles. These results do not immediately necessitate the rejection of the hypothesis that partial loss of *ARID1A* and *ARID1B* alleles confers a growth advantage. In both the CRISPR and shRNA knockdown experiments, only one measurement of growth was examined; it is possible that other advantages could be conferred to these cells, such as enhanced invasion or evasion of apoptosis. It is also possible that because this cell line contains pre-existing mutations in other SWI/SNF subunits\(^5\) and it is from the large intestine lineage, the relationship between ARID1A and ARID1B is different in this cell line. It will be important to study these proteins in a SWI/SNF-wildtype line and perhaps even a non-cancer cell line to better understand the precise interplay between them.

Next, I knocked down ARID1A or ARID1B in the HCT116 parental (*ARID1A* wildtype) cell line via shRNA (Figure 3-5A) to directly compare the effect of the loss of these proteins on cell viability. The growth of the parental line was not affected by ARID1A knockdown (Figure 3-5B), consistent with the reported role for ARID1A as a tumor suppressor\(^6\). Conversely, and consistent with the findings described above, cell growth was significantly impaired by ARID1B loss (Figure 3-5B), suggesting that in this cell line, ARID1B plays a crucial role that is distinct from ARID1A. These results contrast with the finding described in Chapter 2 that ARID1B loss does not affect the growth of ES-2 cells (Figure 2-2B, C), an additional *ARID1A* WT cell line, suggesting that the roles of ARID1A and ARID1B may vary in different cell types.
Figure 3-5. **ARID1A and ARID1B knockdown in HCT116 cells.** A. Western blot confirming ARID1A and ARID1B knockdown following shRNA treatment. B. Colony forming assay showing growth response of the cells to shRNA treatment.
The observations that in \textit{ARID1A} wildtype HCT116 cells, \textit{ARID1B} genetic loss does not affect proliferation (Table 3-1 and Figure 3-2C) and that ARID1B protein knockdown severely impairs cell growth (Figures 3-4 B,C and 3-5B) are not necessarily discordant. In the shRNA-treated cells, the effect of ARID1B loss on the entire cell population is observed. In contrast, in the CRISPR experiments, only the clones that could survive \textit{ARID1B} loss were selected. These clones may have adapted in some way to accommodate ARID1B loss, such as upregulating new isoforms of ARID1A or ARID1B. The observation that only one \textit{ARID1A} WT, \textit{ARID1B} hom clone was generated suggests that loss of ARID1B is a detrimental event for the cells. Together, these two experiments show that loss of ARID1B at both the genetic and the protein levels strongly impairs cell proliferation in HCT116 cells.

The composition of SWI/SNF complexes before and after ARID1A and ARID1B knockdown was investigated in both HCT116 and ES-2 cell lines. Given the differential effect of ARID1A and ARID1B loss on cell viability in these lines, I hypothesized that the loss of these proteins may differently affect the composition of SWI/SNF complexes. An antibody against SMARCC1 was used to pull down SWI/SNF complexes and immunoblot for SWI/SNF subunits showed the makeup of the complex in these cell lines following ARID1A or ARID1B loss. In ES-2 (\textit{ARID1A} wildtype) ovarian cancer cells, loss of ARID1A or ARID1B did not substantially affect the incorporation of other subunits into the complex (Figure 3-6). ARID1A knockdown resulted in slight decrease of ACTL6A, SMARCA4, and SMARCA2 in the complex, while the other subunits were not affected. The only change in SWI/SNF complex composition following ARID1B knockdown in ES-2 cells was that ARID1A incorporation into the complex increased,
likely reflective of ARID1A compensating for ARID1B loss. A higher molecular weight
form of ARID1B was specifically pulled down in the ARID1B knockdown samples,
possibly indicating the presence of an upregulated isoform of ARID1B.

<table>
<thead>
<tr>
<th>ARID1A shRNAs</th>
<th>ARID1B shRNAs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Input</td>
<td>IP</td>
</tr>
<tr>
<td>+ dox</td>
<td>A3</td>
</tr>
<tr>
<td>(induced):</td>
<td>-</td>
</tr>
<tr>
<td>ARID1A</td>
<td></td>
</tr>
<tr>
<td>ARID1B</td>
<td></td>
</tr>
<tr>
<td>SMARCA4</td>
<td></td>
</tr>
<tr>
<td>SMARCA2</td>
<td></td>
</tr>
<tr>
<td>SMARCC1</td>
<td></td>
</tr>
<tr>
<td>SMARCC2</td>
<td></td>
</tr>
<tr>
<td>SMARCB1</td>
<td></td>
</tr>
<tr>
<td>SMARCD1</td>
<td></td>
</tr>
<tr>
<td>SMARCE1</td>
<td></td>
</tr>
<tr>
<td>ACTL6A</td>
<td></td>
</tr>
<tr>
<td>ACTIN</td>
<td></td>
</tr>
</tbody>
</table>

Figure 3-6. Composition of SWI/SNF complexes following ARID1A or ARID1B
knockdown in ES-2 cells. Immunoprecipitation of SWI/SNF complexes from the
nuclear extract of ES-2 cells using a SMARCC1 antibody.
In contrast to the experiments in ES-2 cells and consistent with the CRISPR immunoprecipitation experiments, ARID1A loss in HCT116 (ARID1A wildtype) large intestine carcinoma cells resulted in a reduced incorporation of the core subunit SMARCA4 and a very modest reduction of SMARCA2 into SWI/SNF complexes (Figure 3-7). ACTL6A levels also decreased following ARID1A knockdown. ARID1B knockdown only slightly affected the incorporation of SMARCA4 and did not change the levels of SMARCA2 (Figure 3-7), potentially identifying ARID1A as having a more important role in SWI/SNF complex integrity than ARID1B in this cell line. In this experiment, ARID1B levels appear to increase upon ARID1A knockdown (Figure 3-7), while in a separate experiment, ARID1A knockdown results in a reduction of ARID1B (Figure 3-5). These results highlight the intrinsic variability within this experimental system, and underscore the complicated relationship between these proteins. The reduction of SMARCA4 upon ARID1A loss is difficult to interpret, as this cell line has been reported to contain a SMARCA4 mutation that abrogates its ATPase activity. It is possible that SMARCA4 performs a non-catalytic function in SWI/SNF complexes, and it is also possible that this particular cell line may rely on SMARCA2 as the catalytic ATPase subunit of SWI/SNF. Perhaps SMARCA4 is the first subunit to be lost from the complex in this cell line because it is catalytically inactive and does not serve a core function. As a result of the SMARCA4 mutation, the relationship between ARID1A and ARID1B, as well as other SWI/SNF complex members, may be different in this cell line. These results also highlight the different roles of these proteins in cell lines of varying lineages, as ARID1A knockdown differentially affects complex composition and
ARID1B knockdown has distinct effects on cell proliferation in ES-2 as compared to HCT116 cells.

<table>
<thead>
<tr>
<th>ARID1A shRNAs</th>
<th>ARID1B shRNAs</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Input</strong></td>
<td><strong>IP</strong></td>
</tr>
<tr>
<td>+ dox (induced):</td>
<td>- + + + +</td>
</tr>
</tbody>
</table>

ARID1A

ARID1B

SMARCA4

SMARCA2

SMARCC1

SMARCC2

SMARCB1

SMARCD1

SMARCE1

ACTL6A

ACTIN

Figure 3-7. Composition of SWI/SNF complexes following ARID1A or ARID1B knockdown in HCT116 cells. Immunoprecipitation of SWI/SNF complexes from the nuclear extract of HCT116 cells using a SMARCC1 antibody.
It is also important to note that while detrimental to cell proliferation, ARID1B loss does not affect the composition of SWI/SNF complexes in HCT116 cells. Conversely, loss of ARID1A is tolerated by the cells despite an impairment in complex composition. Perhaps ARID1B has a role outside of SWI/SNF complex composition, such as regulating a crucial gene necessary for HCT116 cell viability, or is involved in a process that is independent of SWI/SNF. Together, these data suggest that in the HCT116 cell line, ARID1A and ARID1B have differential effects on cell growth and complex composition, supporting the hypothesis that these two proteins have unique roles.

**Evaluation of gene expression changes following ARID1A and ARID1B knockdown**

The effect of ARID1A and ARID1B loss on gene expression was assessed via shRNA-mediated knockdown in HCT116 parental (ARID1A WT) cells followed by RNA-seq. Three independent doxycycline-inducible shRNAs targeting either ARID1A or ARID1B were used in this experiment. Following shRNA treatment and induction with doxycycline, RNA and protein were isolated from the cells, and cDNA libraries were prepared for sequencing.

Overall, the changes in gene expression upon ARID1A loss are greater than the changes induced by ARID1B loss (Figure 3-8A). While many of the genes that change following ARID1A or ARID1B knockdown are the same, each regulates a subset of unique genes (Figure 3-8A, asterisks); this cluster of genes corresponds to developmental terms, further supporting the concept of unique roles for these proteins in certain developmental contexts. Of the genes that are downregulated upon ARID1A loss, 10% are also downregulated upon ARID1B loss, and 6% of the genes upregulated by ARID1A
loss are also upregulated when ARID1B is knocked down (Figure 3-8B). This preliminary result supports the initial hypothesis that ARID1A and ARID1B have overlapping roles, but that each also has unique functions.

**Figure 3-8. Gene expression changes following ARID1A or ARID1B knockdown in HCT116 cells.** A. Heat map displaying the ratio of normalized counts. Green indicates genes upregulated upon protein loss and red indicates downregulated genes. * highlights a gene cluster with opposing gene expression changes following ARID1A or ARID1B loss. B. Venn diagram displaying overlap of gene expression changes upon loss of ARID1A or ARID1B.
It is important to acknowledge a caveat with the data that could potentially affect the interpretation of results. Unsupervised hierarchical clustering shows that the samples treated with ARID1A shRNAs, both induced and uninduced, cluster together and separately from the ARID1B shRNA-treated samples (Figure 3-9). This finding is contrary to expectations; I hypothesized that the uninduced samples would cluster together, serving as an internal control for the experiment. Several explanations could explain this unexpected observation in the data. The first and most likely scenario is that the shRNA had some effect on the target gene even in the uninduced samples, causing a small degree of knockdown; i.e. leaky expression of the shRNA construct in the absence of doxycycline. It is also possible that these shRNAs had distinct off target effects that affected both the induced and uninduced samples. Other possibilities include the fact that these shRNAs were from different backbones, the viral integration site of each shRNA may have differed, or that the multiplicity of infection of the viruses delivering the shRNAs may have varied. Despite this possible limitation with the data, I moved forward with the analysis and critically evaluated each conclusion with the known caveats.
Figure 3-9. Correlation of counts across samples. Unsupervised hierarchical clustering with Pearson correlation comparing ARID1A shRNA-treated samples to ARID1B shRNA-treated samples.

In order to reconcile the finding that ARID1B knockdown significantly affected HCT116 cell growth, but ARID1A knockdown did not affect cell growth and had a larger effect on gene expression, I specifically examined the genes that were changed by ARID1B knockdown only. I hypothesized that ARID1B would regulate a subset of key genes that are essential in this cell line. Some of the most interesting genes showed a differential change in expression upon ARID1A and ARID1B knockdown. One example is *NLRP1*, a key mediator of apoptosis and inflammation, which goes up upon ARID1A knockdown and goes down upon ARID1B knockdown, supportive of the reported
opposing roles for ARID1A and ARID1B in cell proliferation\(^1\). An additional
differentially regulated gene is \textit{RNF149}, an E3 ubiquitin ligase\(^8\), which increases
following ARID1A knockdown and decreases following ARID1B knockdown, consistent
with the known association of ARID1B with an E3 ubiquitin ligase\(^9\). This analysis has
identified specific genes that are differently regulated by ARID1A and ARID1B and may
be at the heart of their differential function in this cell line.

Gene set enrichment analysis (GSEA) was performed using the differential
expression values from an average of the three replicates for each sample. Genes
involved in immune response, such as a list of STAT3 targets and interferon responsive
genes, were strongly correlated with genes upregulated upon ARID1A loss, but were not
correlated with genes upregulated following ARID1B loss (Figure 3-10A). Alternatively,
ARID1B loss enriched for an upregulation of gene sets related to cell cycle, including G1
to S transition and cell cycle checkpoints (Figure 3-10B). These gene sets were less
significantly upregulated after ARID1A knockdown, potentially indicating that these
genes were specifically affected by ARID1B loss. Similarly, gene ontology (GO) analysis
on the differentially expressed genes identified many immune response related datasets in
the top ten upregulated pathways upon ARID1A loss (Figure 3-10C), while there were
not any enriched pathways upon ARID1B loss. These results could either indicate that
ARID1A loss has a greater effect on gene expression in HCT116 cells, or that there was a
technical issue with the ARID1B shRNA-treated cells, such as knockdown in the
uninduced samples. Despite this possible confounder, these data suggest that ARID1A
loss affects genes involved in cellular immune response, and that this effect is unique to
ARID1A loss. This finding is consistent with a recent study implicating ARID1A
involvement in cancers driven by inflammation\textsuperscript{10}, further suggesting that ARID1A may play a role in modulating the immune response in an oncogenic setting.
Figure 3-10. GSEA and GO analysis from gene expression data. A. Gene sets enriched upon ARID1A knockdown. ES: enrichment score, NES: normalized enrichment score. B. Gene sets enriched upon ARID1B knockdown. C. Gene ontology terms enriched in genes that are upregulated greater than two-fold upon ARID1A knockdown.
Figure 3-10 (Continued)

A.

<table>
<thead>
<tr>
<th>Statistical measure</th>
<th>STAT3_TARGETS</th>
<th>INTERFERON_RESPONSIVE_GENES</th>
</tr>
</thead>
<tbody>
<tr>
<td>ES ARID1A</td>
<td>0.895</td>
<td>0.871</td>
</tr>
<tr>
<td>NES ARID1A</td>
<td>2.415</td>
<td>2.367</td>
</tr>
<tr>
<td>ES ARID1B</td>
<td>0.360</td>
<td>0.318</td>
</tr>
<tr>
<td>NES ARID1B</td>
<td>1.115</td>
<td>1.005</td>
</tr>
</tbody>
</table>

B.

<table>
<thead>
<tr>
<th>Statistical Measure</th>
<th>G1_S_TRANSITION</th>
<th>CELL_CYCLE_CHECKPOINTS</th>
</tr>
</thead>
<tbody>
<tr>
<td>ES ARID1A</td>
<td>0.602</td>
<td>0.561</td>
</tr>
<tr>
<td>NES ARID1A</td>
<td>1.798</td>
<td>1.701</td>
</tr>
<tr>
<td>ES ARID1B</td>
<td>0.558</td>
<td>0.533</td>
</tr>
<tr>
<td>NES ARID1B</td>
<td>2.000</td>
<td>1.929</td>
</tr>
</tbody>
</table>

C.

<table>
<thead>
<tr>
<th>term_name</th>
<th>term_in Genome</th>
<th>term_in set</th>
<th>p-value</th>
<th>q-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>defense response to other organism</td>
<td>138</td>
<td>20</td>
<td>1.8e-20</td>
<td>1.09e-16</td>
</tr>
<tr>
<td>type I interferon signaling pathway</td>
<td>42</td>
<td>14</td>
<td>4.3e-20</td>
<td>1.09e-16</td>
</tr>
<tr>
<td>cellular response to type I interferon</td>
<td>42</td>
<td>14</td>
<td>4.3e-20</td>
<td>1.09e-16</td>
</tr>
<tr>
<td>response to type I interferon</td>
<td>43</td>
<td>14</td>
<td>6.33e-20</td>
<td>1.2e-16</td>
</tr>
<tr>
<td>defense response to virus</td>
<td>88</td>
<td>16</td>
<td>3.78e-18</td>
<td>5.74e-15</td>
</tr>
<tr>
<td>response to external biotic stimulus</td>
<td>324</td>
<td>23</td>
<td>2.13e-16</td>
<td>2.32e-13</td>
</tr>
<tr>
<td>response to other organism</td>
<td>324</td>
<td>23</td>
<td>2.13e-16</td>
<td>2.32e-13</td>
</tr>
<tr>
<td>response to biotic stimulus</td>
<td>335</td>
<td>23</td>
<td>4.44e-16</td>
<td>4.22e-13</td>
</tr>
<tr>
<td>response to virus</td>
<td>151</td>
<td>17</td>
<td>1.39e-15</td>
<td>1.18e-12</td>
</tr>
<tr>
<td>defense response</td>
<td>677</td>
<td>29</td>
<td>8.31e-15</td>
<td>6.32e-12</td>
</tr>
</tbody>
</table>
To follow up on some specific genes that were affected by ARID1A and ARID1B loss, the expression of these genes was interrogated using RT-qPCR in an independent biological replicate of the HCT116 cells with ARID1A and ARID1B knockdown. Several genes validated the results from the RNA-seq data, reflecting both concordant and discordant changes upon ARID1A and ARID1B loss. For example, the cancer-related genes, *DHRS2*, involved in MDM2-mediated p53 degradation\(^\text{11}\), and *TNS4*, important for cell migration\(^\text{12}\), both decreased upon ARID1A and ARID1B knockdown (Figure 3-11A), supportive of the hypothesis that ARID1A and ARID1B can regulate some of the same genes. In contrast, *MSX1*, a gene important for craniofacial development\(^\text{13}\), increased upon ARID1A knockdown and decreased with ARID1B knockdown, while *RUNX2*, an essential gene for osteoblast differentiation\(^\text{14}\), conversely decreased upon ARID1A knockdown and increased with ARID1B knockdown (Figure 3-11B). These developmental related genes highlight an area where ARID1A and ARID1B may have differential roles. Since these experiments were conducted in a cancer cell line, it will be important to validate these findings in a developmentally relevant, non-cancer cell line system, such as in embryonic stem cells.
Figure 3-11. RT-qPCR validation of genes that changed in RNA-seq data.

A. Expression of genes which changed concordantly upon ARID1A and ARID1B loss. Histogram represents values from RT-qPCR, table displays results from RNA-seq experiment. B. Expression of genes which changed discordantly upon ARID1A and ARID1B loss. Histogram represents values from RT-qPCR, table displays results from RNA-seq experiment.
**Conclusion**

Taken together, these results portray a complicated relationship between the SWI/SNF complex subunits ARID1A and ARID1B. While the synthetic lethality upon loss of both proteins appears to be universal, the precise nature of their interaction is cell context dependent. ARID1A and ARID1B have redundant functions, but these studies have also shown that they retain specific roles depending on cell lineage and mutation background. These experiments also display variability in the expression of the reciprocal subunit following ARID1A or ARID1B knockdown, further suggesting that additional factors within the cells, such as cell cycle state, rate of transcription, or degree of protein knockdown, could affect the way these proteins relate to each other. As a result, variables from both outside and inside the cell could affect ARID1A and ARID1B interaction and function. Regardless, these results support that ARID1A and ARID1B are intricately related, though the precise nature of their relationship remains unclear.

The relationship between ARID1A and ARID1B is reflective of the heterogeneity of SWI/SNF complexes, as multiple variants of these complexes are present in different cell types and developmental stages. These studies have shed light on a complex interplay between two proteins, enhancing our understanding of the biology and furthering the opportunity for therapeutic benefit.

**Methods**

**CRISPR-mediated deletion of ARID1B**

Guide sequences targeting *ARID1B* were designed using publicly available tools\(^4\). These guides were cloned into the PX330 backbone as described\(^4\). Plasmids
containing CRISPR guides were transfected (Turbofect, Thermo FERR0531) into cells of interest along with a GFP expressing plasmid at a ratio of 10:1. To enrich for cells containing the CRISPR plasmids, the top 3-5% GFP expressing cells were FACS sorted. These cells were plated in 96 well plates at a dilution of one cell per well in order to generate single clones. The remaining “bulk” cells were plated together and used for testing of primers.

To screen clones for deletion, cells were trypsinized, pelleted, and resuspended in QuickExtract DNA extraction solution (Epicentre QE09050). Genomic DNA was extracted by incubating the cells at 65°C for 6 minutes and then 98°C for 2 minutes. This genomic DNA was screened for the deletion using primers that amplified two regions: one region that would only amplify if the deletion were present, and a second region that would amplify in cells not containing the deletion. PCR reaction was conducted using HotStarTaq (Qiagen 203443) according to the following conditions: 95°C for 15 minutes; 35 cycles of 95°C for 15 seconds, 60°C for 1 min, 72°C for 1 min; and finally 72°C for 10 minutes. PCR products were run on a 1.3% agarose gel to identify positive clones.

Cell proliferation, protein, and RNA levels were assessed as described in Chapter 2.

**shRNA-mediated knockdown**

HCT116 cells were obtained from Horizon Discovery (HD 104-031 and HD 104-049). shRNAs were lentivirally transduced into cells as described in Chapter 2. shRNA expression was induced with doxycycline treatment. Recognition sequence for ARID1A shRNA A3 is AGGAGCTATCTCAAGATTC, A4 is AGCGAGACACAGCTATTTA, and 2260 is CCTGAGCCTTCAGTCAAGACC. Recognition sequence for ARID1B
shRNA 4205 is GCCGAATTACAAACGCCATAT, 4497 is GCACGCAATGATATGCCTTAT, and 6961 is TTGCTGTCTAGTGCATTCAAA. Co-immunoprecipitation experiments and immunoblots were performed as described in Chapter 2, with the exception that the ARID1A antibodies from Cell Signaling Technologies (12354S), Novus (NB100-55334), and Bethyl (A301-040A) were used.

**RNA-sequencing**

RNA was isolated from the cells as described in Chapter 2, except that following RNA isolation, the samples were further purified using the RNeasy MinElute clean up kit (Qiagen 74204). Quality of RNA was assessed with Agilent High Sensitivity RNA ScreenTape Assay, and all samples passed quality control with a RIN value greater than 8.6. Sequencing libraries were prepared according to the Low Sample Protocol from Illumina (TruSeq RNA Sample Preparation v2 guide) using 4 µg of RNA. Quality of RNA libraries was assessed with the Agilent High Sensitivity DNA ScreenTape Assay. Samples were pooled and sequenced (12 samples in one lane, sequenced twice) on a HiSeq2000 machine at the Center for Cancer Computational Biology (CCCB) at the Dana Farber Cancer Institute.

Data was analyzed in collaboration with Burak Alver, a computational biologist in the Park Lab at the Center for Biomedical Informatics (Harvard Medical School). Briefly, reads were aligned to the reference genome (hg19) using Tophat and the amount of reads between experimental conditions was quantified and compared using Cufflink and Cutdiff software\textsuperscript{16}. Gene set enrichment analyses and gene ontology analyses were conducted on the lists of differentially expressed genes.
RT-qPCR follow up was conducted as described in Chapter 2. Evaluation of protein knockdown was performed as described in Chapter 2.

References


Chapter 4

SMARCB1 loss affects DNA methylation at a subset of upregulated enhancers
Abstract

Rhabdoid tumors, characterized by a sole recurrent mutation in the SWI/SNF core subunit \textit{SMARCB1}, are extremely aggressive malignancies that are also genomics stable. As a result, it is hypothesized that an epigenetic mechanism underlies the oncogenic drive of these cancers. Here, we investigate DNA methylation as a candidate epigenetic process that could promote oncogenesis upon \textit{SMARCB1} mutation. Overall, DNA methylation does not change dramatically when \textit{Smarcb1} is deleted in primary mouse fibroblasts or when SMARCB1 is re-expressed in SMARCB1-deficient human malignant rhabdoid tumor cell lines. SMARCB1 addback does, however, have a subtle, yet specific demethylation effect at enhancers that are upregulated upon re-introduction of this protein. Gene expression at these enhancers also increases upon addback of SMARCB1. These data point to a specific secondary effect of SMARCB1 reintroduction on DNA methylation at a subset of enhancers. These results could further suggest that the SWI/SNF complex plays a role in maintaining DNA methylation at certain genomic loci and also provides a partial explanation for how the mutant complex functions in cancer.
Contributions:

The following individuals contributed to the work described in this chapter: Katherine H. Walsh\textsuperscript{1,2,3,4}, Courtney G. Sansam\textsuperscript{1,2,3}, Xiaofeng Wang\textsuperscript{1,2,3}, Arjen Brinkman\textsuperscript{5}, Chip Stewart\textsuperscript{6}, and Burak Alver\textsuperscript{7}

\textsuperscript{1}Department of Pediatric Oncology, Dana-Farber Cancer Institute, Boston MA 02215, USA
\textsuperscript{2}Division of Hematology/Oncology, Children's Hospital Boston MA, USA
\textsuperscript{3}Department of Pediatrics, Harvard Medical School, Boston MA, USA
\textsuperscript{4}Biological and Biomedical Sciences Program, Harvard Medical School, Boston MA, USA
\textsuperscript{5}Radboud University, Department of Molecular Biology, Nijmegen Centre for Molecular Life Sciences, Faculty of Science, 6500 HB Nijmegen, The Netherlands
\textsuperscript{6}Broad Institute of Harvard and Massachusetts Institute of Technology, Cambridge, Massachusetts, USA
\textsuperscript{7}Center for Biomedical Informatics, Harvard Medical School, Boston MA, USA

K. Walsh designed and performed the SMARCA4 ChIP-seq in MEFs and the SMARCB1 re-expression in rhabdoid tumor cell lines. C. Sansam performed the knockout of Smarcb1 and Smarca4 in MEFs. X. Wang performed ChIP-seq and RNA-seq in rhabdoid tumor cell lines with SMARCB1 re-expression. A. Brinkman analyzed MEF methylation data. C. Stewart analyzed rhabdoid tumor cell line methylation data. B. Alver correlated ChIP-seq and RNA-seq data with rhabdoid tumor cell line methylation data.
Introduction

SMARCB1, a core subunit of the SWI/SNF complex, is a bona fide tumor suppressor in malignant rhabdoid tumors. As detailed in Chapter 1, mutations in SMARCB1 are the sole recurrent event in this otherwise genomically stable, highly malignant cancer of young children. It is unclear how SMARCB1 mutation can lead to the onset and propagation of cancer. One hypothesis to explain oncogenesis in the absence of genomic instability is that an epigenetic mechanism is the driver in these cancers. DNA methylation, a somatically heritable epigenetic process, has been implicated in the progression of several cancers, as oncogenes are often found to be hypomethylated and tumor suppressors are sometimes hypermethylated. We therefore hypothesized that SMARCB1 loss could drive cancer through alteration of these DNA methylation processes.

Results

SMARCA4 does not bind to regions of methylated DNA in MEFs

In order to evaluate changes in DNA methylation following loss of SWI/SNF subunits, we utilized conditional mouse embryonic fibroblasts (MEFs) established in the Roberts lab. These cells are engineered to have loxP sites flanking the gene of interest, and an adenoviral-Cre vector has rapid and highly efficient recombination at this locus. Upon treatment with Cre adenovirus, the genes encoding the core subunits Smarcb1 or Smarca4 were deleted. Genomic DNA was isolated from these cells and was submitted for MethylCap-Seq as a way to measure the changes in DNA methylation following Smarcb1 or Smarca4 loss. In addition, chromatin immunoprecipitation followed by
sequencing (ChIP-seq) with an antibody against SMARCA4 was conducted in the wildtype MEFs as an indicator of the location of SWI/SNF binding in these cells. We hypothesized that if DNA methylation is a contributor to the oncogenesis of SMARCB1-deficient cancers, SWI/SNF complexes would be bound to regions where DNA methylation changed following Smarca4 or Smarcb1 deletion, potentially identifying misregulated tumor suppressors or oncogenes.

The differentially methylated regions (DMRs) following Smarcb1 and Smarca4 loss overlapped with each other, as expected since both genes are core members of the same complex (Figure 4-1A). Surprisingly, however, these DMRs did not overlap with SMARCA4 ChIP-seq peaks, suggesting that SWI/SNF was not bound at any of the regions where DNA methylation changed (Figure 4-1A). Further, SMARCA4 binding peaks did not overlap with the majority of the MethylCap peaks, implying that SWI/SNF may not localize to regions of methylated DNA (Figure 4-1B). These preliminary results point to a potentially mutually exclusive relationship between DNA methylation and SMARCA4, and therefore SWI/SNF, binding. Additional investigation was needed before we could definitively reject our initial hypothesis, and therefore we continued to study this epigenetic mechanism.
Figure 4-1. Differentially methylated regions (DMRs) in MEFs do not correlate with SWI/SNF binding. A. Overlap of DMRs in MEFs following Smarcb1 (Snf5) or Smarca4 (Brg1) knockout correlated with SMARCA4 (BRG1) binding sites. B. Correlation of SMARCA4 (BRG1) binding sites with sites of DNA that are methylated.
**SMARCB1 re-expression results in a decrease of DNA methylation at enhancers in rhabdoid tumor cell lines**

We next sought to evaluate DNA methylation changes upon SWI/SNF perturbation in a complementary and more clinically relevant experimental system. SMARCB1 was re-expressed in SMARCB1-deficient human rhabdoid tumor cell lines, and the DNA methylation level of the genomic DNA from these cells was analyzed on an Infinium Methylation array. Re-expression of SMARCB1 was confirmed by Western blot (Figure 4-2A). The cell lines used for this experiment are: BT12, BT16, G401, TM87, and TTC642. These cell lines originate from different tissues, providing a broad representation of the types of rhabdoid tumors. BT12 and BT16 originate from the brain, TM87 from muscle, G401 from kidney, and TTC642 from an extra-renal rhabdoid tumor. Given our data from the previous experiment in MEFs, we revised our initial hypothesis, expecting that DNA methylation may not play a direct role in the tumorigenic mechanism of SMARCB1 mutation in rhabdoid tumors.

The global DNA methylation levels were compared between samples containing SMARCB1 and samples deficient for SMARCB1. Preliminary analysis showed no change in DNA methylation at the 17,000 genes annotated in the array, but did show a slight decrease in DNA methylation levels following SMARCB1 addback upon examination of all 485,000 probes in the array. Because we saw a change in all the probes as opposed to the annotated genes, we next examined DNA methylation changes at promoters and enhancers. Overall, the level of DNA methylation at promoters is low and there is no change in DNA methylation levels at promoters of SMARCB1-deficient cells as compared to those of re-expressed cells (Figure 4-2B).
Figure 4-2. DNA methylation decreases at enhancers upon SMARCB1 re-expression. A. Western blot confirming addback of SMARCB1 in four of the five cell lines used. B. and C. Difference in methylation beta values between wildtype (WT) cell lines and cell lines with SMARCB1 addback at (B) promoters and (C) enhancers.
In contrast, DNA methylation levels were much higher at enhancers than promoters, and DNA methylation at enhancers decreased in cells with SMARCB1 re-expression (Figure 4-2C). This change was most pronounced in the G401 human rhabdoid tumor cell line (Figure 4-2B and C, top panels). This result suggests that part of the mechanism of oncogenesis following SMARCB1 mutation could be a subtle effect on DNA methylation at enhancers. It is important to note that the MethylCap assay, used for the experiments in MEFs, performs best in regions with high CpG content, and is most sensitive in CpG islands. As a result, this assay may miss some differential methylation in regions such as enhancers, which are not CpG rich. Therefore, the results between the MEFs and the rhabdoid cell lines are not necessarily contradictory; it is possible that the MethylCap assay did not detect these subtle differences in DNA methylation at enhancers in MEFs.

Addback of SMARCB1 in rhabdoid cell lines results in a specific demethylation of a subset of upregulated SWI/SNF-bound enhancers

To further study the change in DNA methylation at enhancers, we specifically investigated the enhancers that change upon SMARCB1 addback. ChIP-seq of the histone marks H3K27ac, to identify enhancers, and H3K4me3, to identify promoters, in G401 and BT16 rhabdoid tumor cell lines with SMARCB1 re-expression had been previously conducted by Xiaofeng Wang, a postdoctoral fellow in the Roberts Lab. Supportive of the results described above, the DNA methylation level at promoters was low, and there was not a change in DNA methylation at upregulated promoters upon SMARCB1 addback (Figure 4-3A, left panel).
Figure 4-3. SMARCB1 addback in rhabdoid cell lines results in specific demethylation of a subset of upregulated SWI/SNF-bound enhancers. A. SMARCB1 re-expression results in a decrease of DNA methylation at some upregulated enhancers. B. Demethylated enhancers (red) are more likely to be bound by SWI/SNF. C. Enhancers that lose DNA methylation are associated with genes that increase expression. Results shown are for G401 cells; similar results were observed in BT16 cells.
Figure 4-3 (Continued)

A. 

![Conservative upregulated promoter and enhancer graphs]

B. 

![G401 enhancer and heatmaps]

Legend:
- **Black**: gain of DNA methylation
- **Red**: loss of DNA methylation
- **Green**: low DNA methylation, no change
- **Blue**: high DNA methylation, no change
- **Magenta**: medium DNA methylation, no change

C. 

![Gene expression and DNA methylation scatter plots]

G401 promoter substitution: 20000, r=0.146

G401 d.TSS>20000, r=0.088

G401 d.TSS>20000, r=0.221
In both cell lines, however, the level of DNA methylation was higher at enhancers and DNA methylation was decreased at enhancers that were upregulated upon SMARCB1 re-expression (Figure 4-3A, right panel). Upregulated enhancers were defined as regions of DNA with a two-fold or greater increase of distal H3K27ac peaks upon SMARCB1 induction.

Next, we sought to determine whether these DNA methylation changes occur at locations of SWI/SNF binding. Xiaofeng Wang had previously conducted ChIP-seq of SWI/SNF complexes in G401 and BT16 rhabdoid tumor cell lines with and without SMARCB1 re-expression, using antibodies against SMARCA4 or SMARCC1, core subunits of the complex. In both G401 and BT16 cells, the enhancers that lose DNA methylation following SMARCB1 induction are more likely to be bound by SWI/SNF than the sites that do not have changes in DNA methylation (Figure 4-3B).

Finally, we asked whether gene expression changes are correlated with the enhancers that are upregulated and demethylated following SMARCB1 induction. We utilized RNA-seq data from G401 and BT16 cells with SMARCB1 re-expression generated by Xiaofeng Wang. The sites that gain H3K27ac (upregulated enhancers) and also lose DNA methylation upon SMARCB1 reintroduction are correlated with an increase in gene expression (Figure 4-3C). These results are supportive of a model where SWI/SNF is responsible for maintaining proper DNA methylation levels at key genes, thereby regulating their transcription. Upon SMARCB1 loss, SWI/SNF complexes can no longer properly modulate DNA methylation, potentially resulting in the aberrant silencing of certain genes. Since the effect on DNA methylation is subtle and only at a subset of upregulated enhancers, it is likely that this change is secondary, and not a direct
result of SMARCB1 loss. Nevertheless, the precise and reproducible change in DNA methylation suggests that the effect is specific and could have biological relevance.

**Conclusion**

These experiments provide evidence that modulation of SWI/SNF complexes could have subtle effects on DNA methylation, thereby affecting gene expression and illuminating part of the mechanism for oncogenesis in a genomically stable background. In this study, SMARCB1 re-expression did not affect DNA methylation at gene bodies or promoters, but did result in a specific loss of DNA methylation at upregulated enhancers. These changes in DNA methylation are specific and are consistent with enhancer and gene expression changes. Taken together, these results suggest that intact SWI/SNF complexes in healthy cells could be involved in the demethylation of DNA at enhancer regions, thereby leading to the upregulation of genes associated with these enhancers. Conversely, loss of this function in SMARCB1-deficient cancers could lead to the aberrant repression of genes important for cancer suppression. Since these changes in DNA methylation occur at so few sites, they are likely to be secondary effects and not a direct consequence of SWI/SNF perturbation. These results identify an indirect relationship between DNA methylation and SWI/SNF, adding to the complicated role of SWI/SNF and its regulation of the epigenetic landscape.
**Methods**

**Cell Culture**

Mouse embryonic fibroblasts were generated as previously described\(^6\). Treatment with Cre adenovirus (Ad5-CMV-Cre) followed by selection in puromycin resulted in deletion of *Smarca4* or *Smarcb1*. G401 (CRL-1441) was purchased from ATCC. TM87, given to the Roberts Lab by Bernard Weissman (University of North Carolina), and TTC642 originated in the laboratory of Timothy Triche (Children’s Hospital of Los Angeles)\(^8\). BT12 and BT16 were a gift to the Roberts lab from Dr. David James (University of California, San Francisco). Cells were transduced with pBabe-puro-SMARCB1 plasmid or pBabe-puro-empty plasmid as control, selected in puromycin, and re-expression of SMARCB1 was confirmed by Western blot, performed as described in Chapter 2. For ChIP-seq and RNA-seq experiments, G401 and BT16 cell lines were transduced with a doxycycline-inducible pInducer21 SMARCB1 re-expression plasmid.

**DNA methylation assays**

Genomic DNA from MEFs with and without *Smarca4* or *Smarcb1* knockout was used for MethylCap-seq, as described\(^7\). Genomic DNA from rhabdoid cell lines with and without SMARCB1 re-expression was analyzed using the Infinium Methylation Array, performed at the Broad Institute in Cambridge MA. The reads were aligned to the reference genome and methylation beta values were compared between samples.

**ChIP-seq and RNA-seq experiments**

Nuclei from wildtype MEFs were isolated and subjected to micrococcal nuclease digestion followed by brief sonication to release nucleosomal fragments. SMARCA4
ChIP was performed on the lysate (SMARCA4 antibody from Millipore: 07-478), and the immunoprecipitated DNA was purified and sequenced\(^9\).

Nuclei from rhabdoid cell lines were isolated as described in Chapter 2. Samples were sonicated to a size of 200 base pairs using Covaris sonicator, and sonication efficiency was confirmed by running samples on an E-gel. Sonicated DNA was immunoprecipitated with antibodies against H3K27ac (Cell Signaling Technology: 8173), H3K4me3 (Millipore: 07-473), SMARCC1 (Santa Cruz: SC-9746), and SMARCA4 (Abcam: ab110641). Libraries for sequencing were prepared with AMPure XP beads (Beckman Coulter: A63881) using the ChIP-seq (TruSeq) Library Amplification protocol (Illumina). Briefly, the sequencing data is aligned to the reference genome and peaks are compared between samples.

RNA-seq of rhabdoid cell lines with and without SMARCB1 re-expression was conducted as described in Chapter 3.

References


Chapter 5

Discussion and future directions
The relationship between ARID1A and ARID1B

Cell context is essential for the precise interaction between ARID1A and ARID1B

The goal of this project was to identify specific vulnerabilities conferred by ARID1A mutation in order to gain a better understanding of SWI/SNF biology and also to identify novel therapeutic opportunities. The data from Project Achilles described in Chapter 2 reveal ARID1B as specifically essential for ARID1A-mutant cancer cell lines. Surprisingly, co-mutation of ARID1A and ARID1B was found in both cell lines and primary cancers, though at least one allele always remained. This observation led to the hypothesis that partial loss of ARID1A or ARID1B alleles confers a growth advantage while loss of all four alleles is lethal. This hypothesis was tested in Chapter 3; the data support the conclusion of synthetic lethality, though a dosage effect was not observed. The evidence from Chapter 3 suggests that the precise interplay between these subunits varies depending on cell context, and therefore partial loss of these proteins may not result in the same effect in all cell types. The relationship between ARID1A and ARID1B is quite complicated, and is reflective of the degree of intricacy in the combinatorial assembly of SWI/SNF subunits.

While the synthetic lethal relationship appears to be universal across lineages, the differential susceptibility of cell lines to ARID1A and ARID1B loss strongly suggests that the roles of ARID1A and ARID1B are cell context dependent. As detailed in Chapter 1, unique variants of SWI/SNF complexes are present at different stages of development\textsuperscript{1–3}, indicating that the composition of these complexes changes depending on developmental state and lineage. ARID1A and ARID1B may similarly change throughout development. Additionally, the mutational landscape of each cell line,
especially the mutation status of genes encoding other SWI/SNF subunits, could affect the roles of, and interplay between, ARID1A and ARID1B. For example, the HCT116 cell line has been reported to contain a point mutation in SMARCA4 that abrogates its ATPase activity and creates a dependency on SMARCA2. The altered dynamic between SWI/SNF subunits could impact the functions of ARID1A and ARID1B in this cell type. Taken together, these data underscore the importance of investigating multiple cell types in order to draw general conclusions regarding SWI/SNF composition and function. The configuration of these complexes is not universal across cell lines, but rather is highly heterogeneous, providing functional specificity to certain lineages and developmental states.

Multiple isoforms of both ARID1A and ARID1B have been reported, and these proteins consistently appear as multiple bands on Western blots. As a result, certain isoforms may be more predominant in specific tissues, and other isoforms may be unique to cell lineages or developmental states. For example, higher molecular weight isoforms of both ARID1A and ARID1B are present in embryos, and their expression is not maintained at the same level throughout development. The possibility of alternatively spliced forms of these proteins adds a layer of complexity to fully understanding their roles. The specific function of each isoform must be considered when assessing ARID1A and ARID1B across cell types.

Even though partial loss of ARID1A or ARID1B in HCT116 isogenic cell lines does not confer a growth advantage, data from multiple cell lines would be needed to reject the proposed dosage hypothesis. Unlike the ovarian cancer cell line ES-2, this cell line exhibits sensitivity to ARID1B loss. Additionally, this cell line is microsatellite
unstable and has mutations in several genes encoding SWI/SNF subunits, such as \textit{SMARCA4}. These results suggest that ARID1A and ARID1B may behave differently across cell lines, and that the dosage relationship may not exist in this particular cell line. In order to thoroughly investigate whether partial loss of ARID1A and ARID1B confers a growth advantage, several cell lines of multiple lineages must be surveyed.

\textit{ARID1A and ARID1B share overlapping roles but also retain unique functions}

An alternate hypothesis to explain the observations of both synthetic lethality and co-mutation of \textit{ARID1A} and \textit{ARID1B} is that these proteins serve redundant functions. In this scenario, mutation of either \textit{ARID1A} or \textit{ARID1B} would result in the same growth advantage. These proteins would be able to substitute for each other in SWI/SNF complexes, and loss of both would be incompatible with cell viability. Some reports in the literature, however, suggest that these proteins do have unique functions, for example in cell cycle progression\textsuperscript{5,7}. It is possible, given the contextual variability of these proteins, that ARID1A and ARID1B can sometimes substitute for each other, but in other situations, these proteins play specific roles.

Upon examination of gene expression changes following ARID1A loss as compared to ARID1B loss, it is evident that depletion of these proteins results in both coordinate and discordant changes in gene expression (Chapter 3). For example, immune response genes were uniquely upregulated following ARID1A knockdown, consistent with a recent report associating \textit{ARID1A} mutation with pro-tumorigenic inflammation\textsuperscript{8}. ARID1B loss alternatively resulted in a significant upregulation of gene sets related to cell cycle progression. Biochemical and cell growth data further support unique roles for
ARID1A and ARID1B; in HCT116 cells, loss of these proteins has differential effects on both SWI/SNF complex composition and proliferation. Because these experiments were conducted in HCT116 cells, the observed changes may not be universal and only apply to this cell line. Further studies are necessary to confirm and extend these findings to other cell types and developmental states. These results are suggestive of a model where ARID1A and ARID1B can substitute for each other in some conditions, but also have specific functions that are unique to each protein.

ARID1A and ARID1B may bind to each other in certain contexts

A third possibility to account for synthetic lethality and co-mutation is that ARID1A and ARID1B, previously thought to be mutually exclusive⁹, could in fact bind together, either in SWI/SNF complexes or in a SWI/SNF-independent context. It is possible that in this scenario, SWI/SNF complexes could sustain loss of either ARID1A or ARID1B, but loss of both would destabilize the complex. Co-mutation of ARID1A or ARID1B, as long as one of the proteins remains intact, may contribute to the aberrant activity of residual SWI/SNF complexes, such as targeting the complexes to improper loci. Preliminary data from rhabdoid, HCT116, and ES-2 cell lines indicate that ARID1A can pull down a higher molecular weight form of ARID1B (Figure 5-1). It is possible that ARID1A could specifically interact with a larger isoform or modified version of ARID1B in some cell types. These ARID1A- and ARID1B-containing complexes may only represent a small minority of SWI/SNF complexes, but it is conceivable that they may exist and be essential in specific cellular contexts. Conversely, ARID1A-ARID1B dimers
could have a function outside of SWI/SNF that renders them indispensable in certain settings.

Figure 5-1. ARID1A IP pulls down a higher molecular weight form of ARID1B.

SMARCB1 expression in rhabdoid cell lines BT16 (A) and TM87 (B) was induced with doxycycline. Nuclear extract was immunoprecipitated with indicated antibodies.

Experiments (A) and (B) were performed by Xiaofeng Wang. C. Nuclear extract from ES-2 and HCT116 cell lines was immunoprecipitated with an antibody against ARID1A in the presence and absence of benzonase.
*Implications for therapy*

The identification of synthetic lethality between ARID1A and ARID1B may present therapeutic opportunity, as *ARID1A* is frequently mutated in many cancers\textsuperscript{10}. While neither ARID1A nor ARID1B have traditionally targetable regions, such as a catalytic site, other potential avenues for therapy exist. One possible therapeutic approach would be to disrupt the interaction between ARID1B and other SWI/SNF subunits in an *ARID1A*-mutant setting, thereby destabilizing the complex and killing the cells. An analogous strategy was shown to be successful in a recent report where stapled peptides successfully disrupted the protein-protein interaction between EZH2 and the Polycomb PRC2 chromatin modifying complex\textsuperscript{11}. ARID1B can also assemble into an E3 ubiquitin ligase that mediates the monoubiquitination of histone H2B\textsuperscript{12}. Targeting the E3 ligase-associated activity of ARID1B is an approach worth considering, although it is unknown whether the E3 ligase activity plays a role in synthetic lethality.

One potential concern for targeting ARID1B in the therapy of *ARID1A*-mutant cancers is the finding that *ARID1B* is also mutated in some cancers\textsuperscript{10}, and that ARID1B inhibition could actually promote cancer in some contexts. In general, SWI/SNF-mutant cancer cell lines remain dependent on the absence of the missing subunit. For example, re-expression of SMARCA4 in *SMARCA4*-mutant cancers results in reversion of malignant phenotype\textsuperscript{13,14}. Therefore, it is likely that a cancer formed by an ARID1B inhibitor would be dependent on that inhibitor, and cessation of the inhibitor would result in the resolution of the tumor. Future studies aimed at dissecting the precise nature between ARID1A and ARID1B, and the nuances of their functions based on cell context, will allow for rationalized and more accurate application for therapy.
Ultimately, for cancers driven by mutation of a gene encoding a SWI/SNF subunit, such as ARID1A, the presence of a residual SWI/SNF complex becomes a specific dependency and potential therapeutic target. This emerging theme, described here for ARID1A and ARID1B, may represent a more universal relationship across cancers with other SWI/SNF subunit mutations. Given the large number of cancers harboring SWI/SNF mutations, investigation of the dependency mechanisms and of the potential to target these residual complexes could have broad cancer relevance.

Future directions

These studies point to two essential areas of future investigation, which will help elucidate the precise nature of the function of ARID1A and ARID1B in SWI/SNF complexes. The first area of experimentation is to examine the composition of ARID1A-containing SWI/SNF complexes as compared to ARID1B-containing complexes. The key experiment will be to immunoprecipitate ARID1A or ARID1B and then examine interacting proteins with mass spectrometry (IP-MS). The major technical barrier to this experiment will be to obtain an efficient antibody with high enough specificity to effectively distinguish between these two proteins. An alternate approach would be to tag ARID1A and ARID1B and immunoprecipitate the proteins with an antibody against the tag.

Following IP-MS, the composition of SWI/SNF complexes containing ARID1A and ARID1B will be compared. If ARID1A and ARID1B are redundant and simply replace each other in BAF SWI/SNF complexes, then the composition of SWI/SNF complexes containing ARID1A or ARID1B will be the same. If, however, specific
ARID1A-containing complexes exist that are distinct from complexes containing ARID1B, the IP-MS will identify unique proteins in each condition. It is also possible that novel proteins may associate with SWI/SNF complexes, and that these unique subunits could help identify the functional differences between ARID1A and ARID1B. This IP-MS experiment will also shed light on any non-conventional roles of ARID1A and ARID1B that may exist. It is possible that these proteins function outside of the SWI/SNF complex, and this experiment could identify non-SWI/SNF interacting proteins that are unique to ARID1A or ARID1B. Additionally, it is possible that ARID1A and ARID1B are not mutually exclusive, as is commonly accepted, and that these proteins associate together in variants of SWI/SNF or other complexes. An important complementary experiment will be to perform a sucrose sedimentation assay with antibodies against ARID1A or ARID1B to examine if the interacting proteins form a complex that is 2 MDa, the size of SWI/SNF, or if these proteins come together to form additional complexes. Additionally, if ARID1A and ARID1B are confirmed to bind to each other, it will be important to study the precise domains necessary for this interaction. These experiments will provide crucial information to understanding how ARID1A and ARID1B relate to each other and how they integrate into SWI/SNF, or other, complexes.

The second aspect of this investigation is to study the localization of ARID1A and ARID1B to DNA by chromatin immunoprecipitation followed by sequencing (ChIP-seq) with antibodies against ARID1A and ARID1B. Again, the primary challenge for this experiment is acquiring ChIP-grade antibodies that can distinguish between these two homologous proteins. If these proteins are completely redundant, then it would be expected that the loci bound by both ARID1A and ARID1B would be identical. If instead
they each retain unique roles, the identification of loci differentially bound by ARID1A as compared to ARID1B will help to clarify the specific functions of each protein. Comparison of ARID1A and ARID1B ChIP-seq to ChIP-seq of a core SWI/SNF subunit, such as SMARCC1, will help determine if ARID1A and ARID1B regulate regions of DNA independently of SWI/SNF. It will also be important to correlate the loci bound by ARID1A or ARID1B with the genes changed following knockdown as measured by RNA-seq and described in Chapter 3.

Together, these experiments will help to further understand the complicated relationship between these two variant homologous subunits of the SWI/SNF complex. An important consideration will be the cell type, lineage, and developmental context to use, and the most thorough examination will include a wide survey of many cell contexts. These experiments have the potential to systematically determine whether ARID1A and ARID1B exist in unique complexes, bind to specific loci, and if they have roles outside of SWI/SNF complexes. Finally, an improved understanding of the precise relationship between these frequently mutated genes in cancer will carry implications for improved therapy.

**The role of DNA methylation as an epigenetic mechanism in SMARCB1-deficient cancers**

**Crucial role for enhancers in rhabdoid tumors**

DNA methylation was investigated as the putative epigenetic mechanism behind genomically stable rhabdoid tumors. The data described in Chapter 4 reveal that re-expression of SMARCB1 in human rhabdoid tumor cell lines leads to a decrease in DNA
methylation at a small subset of upregulated enhancers. In addition, gene expression increases concomitantly with the decrease of DNA methylation. These results point to enhancers as a region of DNA potentially involved in the progression of SMARCB1-deficient rhabdoid tumors and help to identify a potential mechanism by which \textit{SMARCB1} mutation affects transcriptional regulation.

Enhancers, defined as elements of DNA that activate the transcription of a gene from a distance\textsuperscript{15}, are increasingly thought to play important roles in development and disease\textsuperscript{16}. Indeed, unpublished data from the Roberts Lab strongly suggest that enhancers are the regions of DNA most affected by SWI/SNF subunit loss or gain, and therefore the sites most regulated by the complex. In rhabdoid cancer cell lines, addback of SMARCB1 leads to a gain of enhancer marks at lineage-defining genes. For example, in the brain-derived BT16 cell line, re-expression of SMARCB1 results in an increase of enhancers at genes important for neural development and WNT signaling. In addition, a recent publication shows that SMARCA4 is necessary for enhancer activation during mesoderm development\textsuperscript{17}. These results provide additional compelling evidence suggesting that a major role for SWI/SNF complexes is to regulate gene expression at enhancers. Modulation of DNA methylation at some of these enhancers may be one of many ways in which SWI/SNF complexes can control transcription.

\textit{DNA methylation change is a secondary effect following SMARCB1 loss}

The subtle, yet specific effects of SMARCB1 addback on DNA methylation indicate that this change is biologically relevant, but is not a direct consequence of SMARCB1 re-expression. Interestingly, the few sites that exhibit DNA methylation loss
also show a coordinate increase in gene expression and enhancer marks. It is possible that
in the process of remodeling chromatin at enhancers by mobilizing nucleosomes or
recruiting additional factors to the DNA, SWI/SNF complexes also affect DNA
methylation levels. This effect on DNA methylation is not the primary mechanism of the
complex, but instead is likely a byproduct of its main function.

Interrogation of several large-scale data sets has further supported the hypothesis
that DNA methylation is a secondary consequence of SWI/SNF perturbation. First, the
data from Project Achilles was used to determine if DNA methyltransferases, such as
DNMT1, DNMT3A, and DNMT3B, were specifically essential for the rhabdoid tumor
cell lines. If proper maintenance of DNA methylation at key genes was essential for
rhabdoid cell viability, we hypothesized that loss of mediators of DNA methylation
would specifically impair the growth of the rhabdoid tumor cells. In a class comparison
of seven rhabdoid tumor cell lines vs. 144 non-rhabdoid lines, none of the DNA
methyltransferases scored as specifically essential, possibly suggesting that maintenance
of precise DNA methylation levels is not a specific vulnerability for rhabdoid cell lines.
Alternatively, it is possible that the methyltransferases serve redundant functions and loss
of any single one is not enough to kill the rhabdoid cells. It is also possible that loss of the
methyltransferases is lethal across many cell lines, and therefore is not a specific
dependency in the rhabdoid cell lines.

We also examined the dataset from the Cancer Target Discovery and
Development Project, CTD\textsuperscript{2}, a small molecule compound screen under the direction of
Stuart Schreiber at the Broad Institute. In this screen, eight rhabdoid tumor cell lines were
profiled and compared to 907 cancer cell lines screened with 496 compounds. Again, we
hypothesized that if preservation of DNA methylation levels at certain genes is crucial for the survival of rhabdoid tumor cell lines, then DNA methylation inhibitors would specifically impair their growth. We found, however, that none of the DNA methylation inhibitors in the screen were effective against rhabdoid tumors, again suggesting that the changes in DNA methylation following SMARCB1 re-expression are most likely secondary.

In addition, we investigated whether SWI/SNF complexes could directly interact with DNA methyltransferases and if this interaction changed upon re-introduction of SMARCB1. Xiaofeng Wang had previously conducted an IP-MS for SMARCC1 in the rhabdoid tumor cells following SMARCB1 re-expression to examine SWI/SNF complex composition in the presence and absence of this core subunit. None of the known DNA methyltransferases were bound to SMARCC1 in either condition, suggesting that SWI/SNF complexes do not physically interact with mediators of DNA methylation.

Finally, we examined the set of genes which lost DNA methylation, gained enhancer marks, and also increased expression in both G401 and BT16 cell lines. Of the 658 upregulated genes in G401 cells and the 433 in BT16, 84 overlapped. This small overlap could point to a specific set of genes regulated by SWI/SNF-mediated DNA methylation. Alternatively, the small amount of overlap could suggest that the genes which exhibit DNA methylation change are lineage specific. A final possibility is that the gene expression change following SMARCB1 addback at this subset of upregulated enhancers is non-specific and not biologically meaningful. Gene ontology analysis of the 84 genes which lose DNA methylation and increase expression in both cell lines only identifies terms involved in cytoskeleton and microtubule formation. A subsequent
comparison examined the enhancer-associated genes that lose DNA methylation to all enhancer-associated genes; significant enrichment of any gene ontology term or pathway was not observed. Further, comparison of the transcription factor motifs between upregulated enhancers that lose DNA methylation vs. all upregulated enhancers did not enrich for any motifs.

Taken together, these results point to a scenario where DNA methylation changes are not directly regulated by SWI/SNF complexes, but rather are a secondary effect of other actions of SWI/SNF complexes. Inhibition of DNA methyltransferases does not serve as a specific vulnerability of SMARCB1-deficient rhabdoid tumor cell lines, and these methyltransferases do not physically associate with SWI/SNF complexes. Finally, examination of the genes that lose DNA methylation upon SMARCB1 addback does not point to any biologically significant pathways or transcription factor motifs. Consequently, DNA methylation changes are likely a secondary effect following SMARCB1 loss and DNA methylation is not a major epigenetic mechanism behind SMARCB1-deficient cancers.

**Future directions**

Several important questions remain as to the nature of this DNA methylation change upon SMARCB1 addback and whether it can be leveraged for therapeutic intervention. One experiment is to further investigate the specific genes displaying altered DNA methylation levels, first through validation with methylation-specific PCR\(^\text{18}\). These genes could then be over-expressed independently and in different combinations in the rhabdoid tumor cell lines to determine if an increase in the expression of these specific
genes affects the growth of the rhabdoid cells. If the change in DNA methylation occurred specifically in order to regulate that gene, then it is possible that over-expression of the gene will impair cell proliferation. If, however, the change in DNA methylation is a non-specific secondary effect, then over-expression of the gene will not affect the growth of the cells.

An additional experiment would be to utilize the MEF system described in Chapter 4. The assay used in this experiment, MethylCap-seq\textsuperscript{19}, was not sensitive for enhancer regions of DNA. Smarcb1 could be knocked out of the MEFs and their genomic DNA could be used for the Infinium Methylation Array that was performed on the rhabdoid cells. DNA methylation would be examined at both up- and downregulated enhancers. If the change in DNA methylation upon Smarcb1 knockout was coordinately opposite to that of SMARCB1 addback, the conclusion can be drawn that these changes are specific.

Another experiment that would further explore whether this change in DNA methylation is truly a secondary effect rather than a primary result of SMARCB1 re-expression would be to treat rhabdoid cell lines with a DNA methylation inhibitor, such as 5-azacytidine, and observe the effects on DNA methylation, gene expression, and enhancer activation. If the changes in DNA methylation are a primary result of SWI/SNF perturbation, then treatment with 5-azacytidine will have a preferential effect on the DNA methylation level of the genes changed by SMARCB1 addback. In the more likely scenario, 5-azacytidine treatment will result in a widespread decrease of DNA methylation regardless of SMARCB1 addback, suggestive of a non-specific secondary effect.
Finally, it will be important to determine whether this change in DNA methylation upon SMARCB1 addback is specific to rhabdoid tumors or if it is more universal across cancers with mutations in genes encoding other SWI/SNF subunits. Similar experiments could be performed following knockdown or re-expression of other SWI/SNF subunits such as SMARCA4, ARID1A, and PBRM1. If the effect on DNA methylation is a non-specific secondary effect of SWI/SNF-mediated transcriptional regulation at different genes, then it is likely that the same subtle effect will be observed following addback of other subunits. If, however, DNA methylation does not change following re-expression of other SWI/SNF subunits, it will indicate that this finding is specific to rhabdoid tumors.

Conclusion

The results described in this dissertation reveal a novel vulnerability of ARID1A-mutant cancers and elucidate partial mechanism by which mutation of the genes encoding subunits of SWI/SNF complexes promotes oncogenesis. The identification of ARID1B as a synthetic lethality in ARID1A-mutant cancers points to a potential new therapeutic approach for the many types of cancers containing ARID1A mutations. This finding also provides insight into the composition of SWI/SNF complexes and how perturbation of one subunit can affect the integration of other subunits into the complex. The results describing DNA methylation changes at upregulated enhancers in response to SMARCB1 re-expression are reflective of an emerging theme that enhancer regions are important for disease progression. In sum, this body of work has generated novel understanding of the
basic function of SWI/SNF complexes; it has also identified new potential avenues of therapy for the many cancers harboring SWI/SNF mutations.

References


